

..... INTRODUCTION BY

DR. ELIYAHU M. GOLDRATT

author of *The Goal*



**THEORY OF
CONSTRAINTS
HANDBOOK**

EDITED BY

James F. Cox III and John G. Schleier, Jr.

Theory of Constraints Handbook

About the Editors

JAMES F. COX III, PhD, CFPIM, CIRM, holds TOCICO certifications in Production and Supply Chain, Performance Measurement, Critical Chain, Strategy and Tactics, and Thinking Processes. He is a JONAH's JONAH, Professor Emeritus, and was the Robert O. Arnold Professor of Business in the Terry College of Business at the University of Georgia. He has conducted numerous academic and practitioner Theory of Constraints workshops and programs on performance measurement, production, supply chains, management skills, project management, and the thinking processes.

Dr. Cox's research has centered on the Theory of Constraints for over 25 years. He has authored or co-authored three books on TOC and almost 100 peer reviewed articles. He was the co-editor of the *APICS Dictionary*, 7th, 8th, 9th, 10th, and 11th editions, and an invited contributor on the topic of *Constraints Management to the Production and Inventory Management Handbook*.

Dr. Cox has been a member of APICS for over 30 years, holding chapter, regional, and national offices. He served on the APICS Board of Directors for four years with two years as VP of Education—Research and served on the APICS Educational and Research Foundation Board of Directors for nine years with four years as President. He was a founding member and elected to the founding Board of Directors of the Theory of Constraints International Certification Organization (TOCICO), a certification organization founded by Dr. Eli Goldratt. He later served as Director of Certification responsible for implementing TOCICO's certification program.

Now retired, **JOHN G. SCHLEIER**, Jr. was President and Chief Operating Officer of the Mortgage Services Division of Alltel, Inc., Executive Vice President of Computer Power, Inc., and

Director of Office Systems and Data Delivery for IBM. In these positions, he directed major software development projects, sales administration, and financial functions. He was also Director of Information Systems for IBM's General Systems Division, where he provided oversight for Development Engineering, Manufacturing, and Headquarters systems. He developed information systems for manufacturing, sales, and IBM strategic planning functions and was winner of an IBM Outstanding Contribution Award. He was a regular lecturer on Strategic Planning at IBM Executive Briefing Centers over a period of 15 years, speaking to CEOs and top executives of major corporations. He frequently took consulting assignments dealing with complex project management issues around the world.

He served on the faculty of The University of Georgia College of Business Administration as IBM Executive in Residence and later as Executive Professor of Management, serving on both the Management Information Systems and Production Operations Management faculties. Mr. Schleier holds TOCICO certification in all disciplines. He co-authored *Managing Operations: A Focus on Excellence*, a college text emphasizing TOC concepts (North River Press, 2003). He also published *Turkey Tales*, a children's book (Tate Publishing, 2010).

Theory of Constraints Handbook

Edited by
James F. Cox III
John G. Schleier, Jr.



New York Chicago San Francisco
Lisbon London Madrid Mexico City
Milan New Delhi San Juan
Seoul Singapore Sydney Toronto

Copyright © 2010 by James F. Cox, III and John G. Schleier, Jr. All rights reserved. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

ISBN: 978-0-07-166555-1

MHID: 0-07-166555-2

The material in this eBook also appears in the print version of this title: ISBN: 978-0-07-166554-4, MHID: 0-07-166554-4

All trademarks are trademarks of their respective owners. Rather than put a trademark symbol after every occurrence of a trademarked name, we use names in an editorial fashion only, and to the benefit of the trademark owner, with no intention of infringement of the trademark. Where such designations appear in this book, they have been printed with initial caps.

McGraw-Hill eBooks are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. To contact a representative please e-mail us at bulksales@mcgraw-hill.com.

Information contained in this work has been obtained by The McGraw-Hill Companies, Inc. ("McGraw-Hill") from sources believed to be reliable. However, neither McGraw-Hill nor its authors guarantee the accuracy or completeness of any information published herein, and neither McGraw-Hill nor its authors shall be responsible for any errors, omissions, or

damages arising out of use of this information. This work is published with the understanding that McGraw-Hill and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

TERMS OF USE

This is a copyrighted work and The McGraw-Hill Companies, Inc. ("McGraw-Hill") and its licensors reserve all rights in and to the work. Use of this work is subject to these terms. Except as permitted under the Copyright Act of 1976 and the right to store and retrieve one copy of the work, you may not decompile, disassemble, reverse engineer, reproduce, modify, create derivative works based upon, transmit, distribute, disseminate, sell, publish or sublicense the work or any part of it without McGraw-Hill's prior consent. You may use the work for your own noncommercial and personal use; any other use of the work is strictly prohibited. Your right to use the work may be terminated if you fail to comply with these terms.

THE WORK IS PROVIDED "AS IS." McGRAW-HILL AND ITS LICENSORS MAKE NO GUARANTEES OR WARRANTIES AS TO THE ACCURACY, ADEQUACY OR COMPLETENESS OF OR RESULTS TO BE OBTAINED FROM USING THE WORK, INCLUDING ANY INFORMATION THAT CAN BE ACCESSED THROUGH THE WORK VIA HYPERLINK OR OTHERWISE, AND EXPRESSLY DISCLAIM ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. McGraw-Hill and its licensors do not warrant or guarantee that the functions contained in the work will meet your requirements or that its operation will be uninterrupted or error free. Neither McGraw-Hill nor its

licensors shall be liable to you or anyone else for any inaccuracy, error or omission, regardless of cause, in the work or for any damages resulting therefrom. McGraw-Hill has no responsibility for the content of any information accessed through the work. Under no circumstances shall McGraw-Hill and/or its licensors be liable for any indirect, incidental, special, punitive, consequential or similar damages that result from the use of or inability to use the work, even if any of them has been advised of the possibility of such damages. This limitation of liability shall apply to any claim or cause whatsoever whether such claim or cause arises in contract, tort or otherwise.

Contents

[Preface](#)

[Acknowledgments](#)

[Section I What Is TOC?](#)

[1 Introduction to TOC—My Perspective](#) *Eliyahu M. Goldratt*

[Focus](#)

[Constraints and Non-Constraints](#)

[Measurements](#)

[The Goal and The Race](#)

[Other Environments](#)

[The Thinking Processes](#)

[The Market Constraint](#)

[Capitalize and Sustain](#)

[Ever Flourishing](#)

[Strategy and Tactic Trees](#)

[New Frontiers](#)

[References](#)

[About the Author](#)

[Section II Critical Chain Project Management](#)

[2 The Problems with Project Management](#) *Ed Walker*

[Introduction](#)

[Purpose and Organization of the Chapter](#)

Traditional Planning and Control Mechanisms in Project Management

Gantt Charts

PERT/CPM in the Single Project Environment

Brief Review of Project Management Literature

Origins of PERT and CPM

Project Failures

Single Project Management Literature

Multiple Project Management Literature

Development of Guidelines

Macro Issues

Micro Issues

A Brief Overview of Critical Chain Project Management

Critical Chain in the Single Project Environment

Brief Review of Critical Chain Literature

Summary and Conclusions

References

About the Author

3 A Critical Chain Project Management Primer

Charlene Spoede Budd and Janice Cerveny

Introduction

Why These Widespread Project-Related Problems Persist

Task Duration Uncertainty

Traditional Survivor Behaviors

Key Elements of Critical Chain

Issues in Creating a Project Plan

Issues in Managing Project Execution

Scheduling a Single Project

Modifying Task Duration Estimates

A Bit of Statistics

Critical Chain Scheduling

Critical Chain Scheduling—Steps 1 through 4

[Merging Paths—Step 5](#)
[Communications—Step 6](#)
[Three Sources of Critical Chain Project Protection](#)
[Scheduling Projects in Multi-Project Environments](#)
[Establishing Project Priorities](#)
[Selecting a Scheduling Resource and](#)
[Establishing Scheduling Buffers](#)
[Project Control: The Power of Buffer Management](#)
[Tracking Buffer Consumption](#)
[Knowing When to Act](#)
[Adjusting Buffers](#)
[Using Buffer Consumption Information to](#)
[Continuously Improve](#)
[Project Budgeting](#)
[Components of a Project Budget](#)
[Assigning Total Project Costs to Project Tasks](#)
[Implementing a New Project Budgeting Process](#)
[Project Reporting](#)
[Internal Reporting](#)
[External Reporting](#)
[Causing the Change: Behavioral Issues,](#)
[Management Tactics, and Implementation](#)
[Managerial Actions to Support Critical Chain](#)
[Project Management](#)
[Importance of Trust](#)
[Implementing a Critical Chain Project](#)
[Management System](#)
[Summary](#)
[References](#)
[About the Authors](#)

4 Getting Durable Results with Critical Chain—A **Field Report** *Realization Technologies, Inc.*

[Background](#)
[Purpose and Organization](#)

Recap of Critical Chain

Rule 1 Pipelining: Limit the Number of Projects in Execution at One Time

Rule 2 Buffering: Discard Local Schedules and Measurements, and Use Aggregate Buffers

Rule 3 Buffer Management: Use Buffers to Measure Execution, and Drive Execution Priorities and Managerial Interventions

Practical Challenges in Implementing Critical Chain

Challenge 1: Gaining Managerial Commitment for Implementing the Three Rules

Challenge 2: Translating Concepts into Practical Procedures and Instructions

Challenge 3: Sustaining the Critical Chain Rules and Results

Step-By-Step Process for Implementing Critical Chain

Step 1: Achieve Management Buy-In

Step 2: Reduce WIP and Implement "Full Kitting"

Step 3: Build Buffered Project Plans

Step 4: Establish Task Management

Step 5: Implement Surrounding Processes

Step 6: Identify Opportunities for Continuous Improvement (POOGI)

Step 7: (When Applicable) Use Superior Delivery as a Competitive Advantage to Win More Business

Lessons Learned

Performance Gains Come from Managing Differently, Not Better Planning and Visibility

Implement All of the Three Rules

Top Managers Must Play an Active Role

Actively Manage the Buffers

Frequently Asked Questions

Can Critical Chain be implemented without basic project management in place first?

[Should a pilot be run before a full rollout of Critical Chain?](#)

[What about cultural and behavioral changes?](#)

[What is the role of software in Critical Chain?](#)

[Is a Project Management Office \(PMO\) needed with Critical Chain?](#)

[How is non-project work handled with Critical Chain?](#)

[Should the scope of a Critical Chain implementation include vendors and subcontractors?](#)

[How does Critical Chain improve quality?](#)

[Critical Chain seems to be all about timelines; what about controlling costs?](#)

[Do we need project-level budgets in multi-project operations?](#)

[Does Critical Chain work with Earned Value Reporting?](#)

[How does Critical Chain work with Lean?](#)

[What are the likely causes of failure in implementing Critical Chain?](#)

[Summary](#)

[References](#)

[About the Author](#)

5 Making Change Stick *Rob Newbold*

[Introduction](#)

[The Uptake Problem](#)

[No Urgency to Change](#)

[The Silver Bullet](#)

[Negative Branches](#)

[Root Causes](#)

[The Cycle of Results \(CORE\)](#)

[Basic Principles](#)

[Simple Example: Cleaning the Room](#)

[Simple Example: TOC Practitioners Group](#)

[Other Processes](#)
[Implementation Planning](#)
[Planning with the Cycle of Results](#)
[Traps](#)
[Summary](#)
[References](#)
[About the Author](#)

6 Project Management in a Lean World— Translating Lean Six Sigma (LSS) into the Project Environment *AGI-Goldratt Institute*

[Introduction: It's a Lean World](#)
[What Is the Project Environment's Point of View to
Being Leaned?](#)
[Project Environment System of Systems](#)
[What Do We Improve?](#)
[Translating Lean into the Project System of
Systems for Improvement](#)
[Addressing the Disconnects in Lean Techniques
for Project Environments](#)
[The Five Principles of Lean Applied to the Project
Environment](#)
[Specifying Value](#)
[Identify Steps in the Value Stream](#)
[Make Value-Creating Steps Flow towards the
Customer](#)
[Let Customers Pull Value from the Next
Upstream Activity](#)
[Pursuing Perfection](#)
[Leaning Traditional Project Management](#)
[References](#)
[About the Author](#)

Section III Drum-Buffer-Rope, Buffer Management and Distribution

7 A Review of Literature on Drum-Buffer-Rope, Buffer Management and Distribution *John H. Blackstone Jr.*

Introduction

Literature on Precursors of TOC and DBR

Historical Developments Preceding TOC

Derivation of DBR Using the Five Focusing Steps

Literature on DBR Scheduling

Overviews

Applying DBR to Different Types of Facilities:

VATI Analysis

Special Cases

Free Goods

What if the Market Is the Constraint?

Re-Entrant Flows

Recoverable Manufacturing and

Remanufacturing

Buffer Management Literature

Buffer Sizing

Buffer Sizing and Lead Time

TOC and Distribution

Supply Chain Management

Service Environment

TOC and Other Modern Philosophies

Problems with DBR

Floating or Multiple Bottlenecks

Summary and Conclusions

References

About the Author

8 DBR, Buffer Management, and VATI Flow Classification *Mokshagundam (Shri) Srikanth*

Introduction

Managing Flow—Planning and DBR

[The Need for a Focus on Flow](#)
[Ford and Toyota Production Systems—A New Perspective](#)
[Production Operations and the Five Focusing Steps of TOC](#)
[Characteristics of Production Operations](#)
[Applying the Five Focusing Steps to Production Operations](#)
[The DBR System](#)
[The Drum](#)
[The Buffer](#)
[The Rope](#)
[Managing Flow with DBR—An Example](#)
[Managing Flow—Controlling Execution and Buffer Management](#)
[The Need for Control and the Need for Corrective Actions](#)
[Understanding Buffers: The Buffer as the Source of Information for Controlling Execution](#)
[Buffer Management—The Process](#)
[Complex Production Environments and a Classification Scheme](#)
[The Fundamental Elements of the Classification Scheme](#)
[V, A, T, and I Flows—Descriptions and Examples](#)
[V-Plants](#)
[DBR in V-Plants](#)
[A-Plants](#)
[DBR in A-Plants](#)
[T-Plants](#)
[DBR in T-Plants](#)
[I-Plants](#)
[DBR in I-Plants](#)
[Summary](#)
[References](#)
[About the Author](#)

9 From DBR to Simplified-DBR for Make-to-Order

Eli Schragenheim

Introduction

A Historical Background and Perspective

Three Views on Operations Planning and Execution

The Five-Focusing Steps (5FS)

The Critical Distinction between Planning and

Execution

Concentrating on the Flow

Challenging the Traditional DBR Methodology

What Should the Strategic Constraint Be?

How Is the Planning and Execution Viewpoint

Addressing the Issue of Scheduling and Buffering the CCR?

How Does Refraining From a Detailed Schedule of the CCR Affect the Execution?

What Does the Emphasis on Flow Add to the Challenge to Traditional DBR?

Outlining the Direction of the Solution

The Main Ingredients of the Solution

The Time Buffer

Load Control

Determining the Safe Dates

Capacity Reservation

Buffer Management

Short-Term Planned Load

The Notion of “Slack”

Where S-DBR Fits Nicely

The Cases Where S-DBR Does Not Fit

Implementation Issues and Processes

Looking Ahead to MTS

Suggested Reading

References

About the Author

10 Managing Make-to-Stock and the Concept of Make-to-Availability *Eli Schragenheim*

Introduction

Why Is a Special Methodology for MTS Required?

The Current Confusion in Managing Stock

The Common Misunderstanding of Forecasts

The Current Undesirable Effects in MTS

What to Do? The Direction of the Solution

The Basic Principle of Flow

From MTS to MTA

Determining the Appropriate Inventory

Buffer Management in MTA

Generating Production Orders and the State of

Capacity

Peak and Off-Peak Behaviors

Monitoring the Target Level Size—Dynamic Buffer

Management

Too Much Green—the Target Is Too High

Too Much Red—the Target Is Too Low

Discussion: Issues with DBM and By How Much

to Increase/Decrease the Targets

The Role of Protective Capacity and the Usefulness

of Maintaining a Capacity Buffer

The Process of Ongoing Improvement (POOGI)

Generic Issues in MTA

MTA for Components

Which Items Fit MTA and Which Fit MTO?

Vendor-Managed Inventory (VMI)

Mixed (MTA and MTO) Environments

Dealing with Seasonality

Problematic Environments for MTA

MTS That Is Not MTA

Implementation Issues

Moving from MTS or MTO to MTA

Software Considerations

[References](#)
[Suggested Reading](#)
[About the Author](#)

11 Supply Chain Management *Amir Schragenheim*

[Introduction: The Current Practice of Managing Supply Chains](#)

[Problems with the Current System](#)

[The Natural Tendency for Push Behavior](#)

[Why Is It Impossible to Find a Good Forecasting Model?](#)

[The TOC Way—The Distribution/Replenishment Solution](#)

[Aggregate Stock at the Highest Level in the Supply Chain:](#)

[The Plant/Central Warehouse \(PWH/CWH\)](#)

[Determine Stock Buffer Sizes for All Chain Locations Based on Demand, Supply, and Replenishment Lead Time](#)

[Increase the Frequency of Replenishment](#)
[Manage the Flow of Inventories Using Buffers and Buffer Penetration](#)

[Use Dynamic Buffer Management](#)
[Set Manufacturing Priorities According to Urgency in the PWH Stock Buffers](#)

[Why Does a Pull Supply Chain Work Better?](#)
[Some of the Finer Points in Managing the TOC Distribution/Replenishment Solution](#)

[Managing Product Portfolios](#)
[Rules for Setting up Initial Buffer Sizes](#)
[Managing Seasonality in the TOC](#)

[Distribution/Replenishment Model](#)
[Known Patterns for Sudden Changes in Consumption](#)

[Two Different Changes](#)

[Resolving the Forecasting versus DBM Dilemma to Provide Excellent Consumption before, during, and after an SDC](#)

[Identifying When an SDC Is Meaningful](#)

[Handling of an SDC](#)

[Implementing the TOC Distribution/Replenishment Model—How Can Software Help and Is It Really Needed?](#)

[Testing the Solution on a Smaller Scale](#)

[Simulation](#)

[Pilot Project](#)

[Managing the TOC Buy-in Process](#)

[Actual Results of the TOC Distribution/Replenishment Solution](#)

[Summary](#)

[References](#)

[Recommended Reading](#)

[About the Author](#)

[12 Integrated Supply Chain](#) *Chad Smith and Carol Ptak*

[Introduction](#)

[Identifying the Real Problem—Rethinking the Scope of Supply Chain Management](#)

[A Brief History of MRP](#)

[Can MRP Meet Today's Challenge?](#)

[The MRP Conflict Today](#)

[The MRP Compromises](#)

[Actively Synchronized Replenishment—the Way Out of MRP Compromises](#)

[1. Strategic Inventory Positioning](#)

[2. Dynamic Buffer Level Profiling and Maintenance](#)

[3. Dynamic Buffers](#)

[4. Pull-Based Demand Generation](#)

[5. Highly Visible and Collaborative Execution](#)

[Case Studies](#)

[Case Study 1: Oregon Freeze Dry](#)

[Case Study 2: LeTourneau Technologies, Inc.](#)

[Summary](#)

[References](#)

[About the Authors](#)

Section IV Performance Measures

13 Traditional Measures in Finance and Accounting, Problems, Literature Review, and TOC Measures *Charlene Spoede Budd*

[Introduction](#)

[Traditional Cost Accounting and Business Environment](#)

[Development of Cost Accounting Business Environment, First Half of the 20th Century](#)

[Business Environment, Second Half of the 20th Century](#)

[Accounting's Response to a 20th Century Changing Environment](#)

[Direct or Variable Costing Income Statement](#)

[Activity-Based Cost Accounting](#)

[Balanced Scorecard](#)

[Lean Accounting](#)

[Traditional Budgeting, Capital Budgets, and Control Mechanisms](#)

[TOC Approach to Planning, Control, and Sensitivity Analysis](#)

[Planning](#)

[Throughput Control](#)

[Sensitivity Analysis](#)

[Throughput Accounting Approach to Performance Evaluation](#)

[Possible Explanations for the Lack of TOC Literature in Accounting and Finance](#)

[Future TOC Accounting/Finance Research Needs](#)

[Case Studies and Simulations](#)

[Information and Decision Making](#)

[Summary and Introduction of Remaining Chapters in This Section](#)

[Chapter Summary](#)

[Other Chapters Dealing with Performance Measures](#)

[References](#)

[About the Author](#)

14 Resolving Measurement/Performance Dilemmas *Debra Smith and Jeff Herman*

[Introduction](#)

[Do We Measure Too Much?](#)

[Why Do We Have Measurements?](#)

[Global Metrics](#)

[The Constraint Is the Primary Relevant Factor](#)

[Profit Maximizing in TOC](#)

[Local Metrics](#)

[Metric 1: Reliability](#)

[Metric 2: Stability](#)

[Metric 3: Speed/Velocity](#)

[Metric 4: Strategic Contribution](#)

[Metric 5: Local Operating Expense](#)

[Metric 6: Local Improvement/Waste](#)

[Feedback and Accountability Systems](#)

[So, How Is the Operational System Performing?](#)

[Focusing on Improvement](#)

[A Case Study](#)

[Summary](#)

[References](#)

[About the Authors](#)

15 Continuous Improvement and Auditing *Dr. Alan Barnard*

Introduction

The Goal—Achieving Continuous or Ongoing Improvement

Purpose and Organization of This Chapter

Key Concepts and Definitions

A Historical Perspective—Standing on the Shoulders of Giants

Why Change?

Introduction

The Improvement Gap and Challenges

The Types of Management Mistakes When under Pressure to Change

The Extent and Consequences of the Failure Rate of Change

The Vicious Cycle Related to the High Failure Rate of Change

Summary of Why Change?

What to Change?

Introduction

Finding the Core Conflicts within Continuous Improvement and Auditing

Finding a Simple and Systematic Way to Break Conflicts

Identifying Limiting versus Enabling Paradigms in Continuous Improvement

Summary of *What to Change*

To What to Change?

Introduction

Criteria to Evaluate a New Solution

Direction of Solution to Breaking the Continuous Improvement Conflicts

Lessons from CI Methods Developed by Ford and Ohno and Other Giants

[Incentives](#) [Importance \(and Risks\) of Measurements and](#)
[Ensuring the New Direction Addresses All Major](#)
[UDEs](#) [Potential Negative Branches and How to Prevent](#)
[Them](#) [Summary of “What to Change to?”](#)
[How to Cause the Change?](#)
[Typical Implementation Obstacles and How to](#)
[Overcome These](#) [Using TOC to Focus and Accelerate Lean and Six](#)
[Sigma Initiatives](#) [Using TOC’s S&T as a CI and Auditing Tool](#)
[Summary of How to Cause the Change](#)
[Summary of Continuous Improvement and Auditing](#)
[the TOC Way.](#) [References](#)
[About the Author](#)
[Appendix A—Continuous Improvement Opportunity](#)
[Templates](#)

16 Holistic TOC Implementation Case Studies *Dr. Alan Barnard and Raimond E. Immelman*

[Introduction](#)
[Historical Perspective to Holistic TOC](#)
[Implementations](#) [The Goldratt Satellite Program](#)
[The X-Y Syndrome of Local TOC](#)
[Implementations](#) [The “4 × 4”—First Attempt at a Process to](#)
[Launch a Holistic TOC Implementation](#) [The Viable Vision Initiative](#)
[Using TOC’s Strategy and Tactic Tree to Guide](#)
[Holistic Implementations](#)

[Catering for Differences within the Private and Public Sector](#)

[Holistic Implementation of TOC in the Public Sector](#)

[Background](#)

[Designing the Five-Day TOC Workshop and Implementation Process](#)

[Proposed Changes to the Traditional TOC TP Analysis Roadmap](#)

[Detailed Case Study: Analysis on Solid Waste Management in City A](#)

[Current Status of Pilot Projects \(by the End of 2009\).](#)

[Future Application of TOC within the Public Sector](#)

[Specific Lessons Learned from All the Public Sector Pilots](#)

[Future Research](#)

[Holistic TOC Implementation in the Private Sector](#)

[The Birth of First Solar Inc.](#)

[Theory of Constraints Contribution to First Solar's Success](#)

[Building the Foundation](#)

[Unbolting the Existing Systems and Measures](#)

[Building on Early Success](#)

[Implementing the Proven TOC Toolset](#)

[The Role of TOC's "Thinking Processes" at First Solar](#)

[What Has Made TOC Work at First Solar?](#)

[Recommendations and Summary](#)

[Recommended Good Practices for Implementing TOC Holistically](#)

[Summary](#)

[References](#)

[About the Authors](#)

Section V Strategy, Marketing, and Sales

17 Traditional Strategy Models and Theory of Constraints *Marjorie J. Cooper*

Introduction

What Is a Business Strategy?

Factors That Comprise Strategy

Criteria for a Good Strategy

Theories of Business Strategy

Ansoff's Matrix of Four Strategies

Porter's List

The Resource-Based View

Learning/Emergent Strategies

A Summary of Schools of Strategy

Marketing and Strategy

What Is Marketing Strategy?

Sales and Strategy

Challenges for Strategy and Execution

Inadequate Planning

Inability to Analyze the System

No Theory of Implementation

Conflicts within the System

Conflicting Standards of Performance

Dysfunctional Compensation and Reward

Policies

TOC Contributions

Future Research Opportunities

References

About the Author

18 Theory of Constraints Strategy *Gerald Kendall*

Introduction—What Differentiates a TOC Strategy?

Chapter Overview

Definitions and Foundations of TOC Strategy

Three Goals or Necessary Conditions of Any

Strategy

[The Five Focusing Steps](#)
[Example—The Five Focusing Steps](#)
[The Role of Throughput Accounting and Other Metrics in Strategy.](#)

[Overview of TOC Strategy Applications in Manufacturing, Projects, and Consumer Goods Distribution/Retail Organizations](#)

[Introduction to Strategy Applications](#)

[Generic Content of S&T Structures](#)

[Manufacturing](#)

[Projects](#)

[Distribution/Retail](#)

[Six Ways That the Holistic Distribution System Increases Throughput](#)

[Four Generic Prerequisites/“Injections” for a Lasting Competitive Edge](#)

[INJ. 1: Increase Customer Perception of Value that Competitors Have Difficulty Copying](#)

[INJ. 2: Implement Practical Segmentation](#)

[INJ. 3: Identify and Build the Decisive Competitive Edge Factor](#)

[INJ. 4: Strategic Segmentation](#)

[Desirable Effects of a Good Strategy.](#)

[Two Forms of Strategy and Tactics—TP and S&T](#)

[Trees](#)

[Integrating Other Methodologies Such as Lean and Six Sigma](#)

[Dealing with Human Behavior in a Strategy.](#)

[Summary.](#)

[References](#)

[About the Author](#)

[19 Strategy H. William Dettmer](#)

[The Popular Conception of Strategy.](#)

[The System Concept](#)

[A Vertical Hierarchy](#)
[A Common Denominator](#)
[A Whole-System View](#)
[The OODA Loop](#)
[Strategy as a Journey](#)
[Orientation and Observation](#)
[Decision and Action](#)
[“Pro-Acting” Rather than Reacting](#)
[Fast OODA Loop Cycles](#)
[Summarizing Boyd](#)
[The Logical Thinking Process](#)
[The Intermediate Objectives Map](#)
[Constraint Management Model: A Synthesis of TOC
and the OODA Loop](#)
[The Role of the LTP in the CMM](#)
[What about Steps 6 and 7?](#)
[Summary and Conclusion](#)
[References](#)
[About the Author](#)

20 The Layers of Resistance—The Buy-In Process According to TOC *Efrat Goldratt-Ashlag*

[Introduction](#)
[The Layers of Resistance to Change](#)
[Disagreement on the Problem](#)
[Layer 0. There is no problem](#)
[Layer 1. Disagreeing on the problem](#)
[Layer 2. The problem is out of my control](#)
[Disagreement on the Solution](#)
[Layer 3. Disagreeing on the direction for the
solution](#)
[Layer 4. Disagreeing on the details of the
solution](#)
[Layer 5. “Yes, but...” The solution has negative
ramifications](#)

[Disagreement on the Implementation](#)

[Layer 6: Yes, but... we can't implement the solution](#)

[Layer 7: Disagreement on the details of the implementation](#)

[Layer 8: You know the solution holds risk](#)

[Layer 9: "I don't think so"—Social and psychological barriers](#)

[Sense of Ownership: The Key to True Buy-In](#)

[Bottom Line](#)

[References](#)

[About the Author](#)

21 Less Is More—Applying the Flow Concepts to Sales *Mauricio Herman and Rami Goldratt*

[Introduction](#)

[Improving Flow](#)

[Preventing Overproduction](#)

[Local Efficiencies Must Be Abolished](#)

[A Focusing Process Must Be in Place](#)

[Summary](#)

[Addendum](#)

[References](#)

[About the Authors](#)

22 Mafia Offers: Dealing With a Market Constraint *Dr. Lisa Lang*

[Introduction: What Is a Mafia Offer?](#)

[Do You Have a Market Constraint?](#)

[Developing a Mafia Offer](#)

[Custom Label Printer—An Example](#)

[The Test—Is It a Mafia Offer?](#)

[What Did It Take to Make the Offer?](#)

[A Mafia Offer Is NOT](#)

[Where to Start?](#)

[Sustaining the Advantage and the Offer](#)

[It's a Business Deal](#)

[The Psychology of Delivering a Mafia Offer](#)

[Agree on the Problem](#)

[Agree on the Direction of the Solution](#)

[Agree the Solution Solves the Problem](#)

[Agree on the Problem](#)

[Agree on the Direction of the Solution](#)

[Agree Our Solution Solves Their Problem](#)

[Close](#)

[For Whom Can You Develop Offers?](#)

[Can You Create a Mafia Offer?](#)

[The Templates](#)

[Vendor Managed Inventory](#)

[Reliable Rapid Response](#)

[Consumer Goods](#)

[Projects](#)

[Pay Per Click](#)

[Gain Sharing \(My Mafia Offer\)](#)

[Summary](#)

[References](#)

[About the Author](#)

Section VI Thinking Processes

23 The TOC Thinking Processes *Victoria J. Mabin
and John Davies*

[Introduction](#)

[Preface to Chapter](#)

[Purpose of the Chapter](#)

[Outline of Chapter](#)

[The Nature, Development, and Use of the TOC TP](#)

[Overview of TP and Their History and](#)

[Development](#)

[The TP Tools](#)

[The TOC TP Literature](#)

[The Nature of Other Approaches to Problem-Solving and Decision Making](#)

[The Relationship of Problem-Solving Methods to Problem-Solving Activity](#)

[Unstructured Approaches—Management on the Hoof](#)

[Formal or Structured Approaches](#)

[Lessons for TOC from the Literature](#)

[Issues Emerging from the TOC Literature](#)

[The Nature of the TOC Literature Vis-à-Vis Other Literatures](#)

[Suggested Topics for a Self-Audit of TOC](#)

[The Nature and Use of the TOC Thinking Processes Revisited](#)

[Understanding the Relationship of the TOC TP to Problem-Solving Activity](#)

[The Philosophical Basis of the TOC TP](#)

[Summary Insights from Classificatory Mapping of the TOC TP](#)

[Summary](#)

[What Has Been Covered in This Chapter](#)

[Findings and Recommendations](#)

[Links to Other Chapters in the TP Section](#)

[References](#)

[About the Authors](#)

24 Daily Management with TOC *Oded Cohen*

[Introduction—Purpose of the Chapter](#)

[Solving Daily Problems](#)

[Problem Investigation and Solution](#)

[Development—the Cloud](#)

[Inner Dilemmas](#)

[Day-to-Day Conflicts](#)

[Reducing Fire Fighting](#)
[Dealing with the Undesirable Effects \(UDEs\)—](#)
[the UDE Cloud](#)

[Example of a System UDE Cloud-Production](#)
[Example of a System UDE Cloud-Retail](#)
[Addressing Multiple Problems—the Consolidated](#)

[Cloud](#)
[From a Problem to the Solution Implementation](#)
[The TOC Methodology for Problem Solving—the](#)

[U-Shape](#)
[Strengthening the Solutions—Dealing with NBRs](#)
[The Intermediate Objective \(IO\) Map and](#)

[Implementation Plans](#)
[Conclusion—Problem Solving the TOC Way](#)
[References](#)
[About the Author](#)

[25 Thinking Processes Including S&T Trees](#) *Lisa J. Scheinkopf*

[Introduction: Anybody Can Be a Jonah!](#)
[The Basic Building Block—Cause-and-Effect Logic](#)
[Basic Terms and Mapping Protocol](#)
[Tools for Daily Decision Making and Problem Solving](#)
[Negative Branch Reservation \(NBR\)](#)
[Evaporating Cloud \(EC\)](#)
[The Integrated TOC Thinking Processes](#)
[Reinforcing the Mentality of a Scientist—Jonah’s](#)

[Approach](#)
[What to Change?](#)
[Current Reality Tree \(CRT\)](#)
[Evaporating Cloud \(EC\)](#)
[The “Snowflake Method”](#)

[The Bank Case: What to Change, Snowflake](#)
[Approach](#)
[The “Three-Cloud Method”](#)

[To What to Change](#)

[Evaporating Cloud](#)

[Future Reality Tree and Negative Branch](#)

[Reservation](#)

[How to Cause the Change](#)

[Prerequisite Tree](#)

[Transition Tree](#)

[The Strategy & Tactic Tree](#)

[The First Step: The Goal](#)

[Communication, Alignment, and Synchronization](#)

[Implementing an S&T](#)

[Using the TPs to Implement an S&T](#)

[The Knowledge Organizer](#)

[Chapter Wrap-Up](#)

[References](#)

[About the Author](#)

[Appendix B: Categories of Legitimate Reservation¹](#)

26 TOC for Education *Kathy Suerken*

[Why Change?](#)

[What to Change?](#)

[What to Change to?](#)

[How to Cause the Change?](#)

[The Cloud](#)

[The Logic Branch](#)

[The Ambitious Target Tree](#)

[A Process of Ongoing Improvement](#)

[References](#)

[About the Author](#)

27 Theory of Constraints in Prisons *Christina Cheng*

[Introduction](#)

[What To Change?](#)

[Preliminary Study](#)
[Stigmatization](#)
[Negative Peer Pressure](#)
[Importance of Face](#)
[What to Change to?](#)
[Self-Regulation](#)
[Why TOC?](#)
[How to Effect the Change?](#)
[Marketing](#)
[Course Materials](#)
[Delivery](#)

¹ For Appendices A and C to G see <http://www.mhprofessional.com/TOCHandbook>.

[Results](#)
[Quantitative](#)
[Qualitative](#)
[Follow-on Implementations](#)
[Future Recommendations](#)
[Summary and Conclusion](#)
[About the Author](#)

Section VII TOC in Services

28 Services Management *Boaz Ronen and Shimeon Pass*

[Introduction](#)
[Challenges in Service Management](#)
[Why the Need for Change?](#)
[Survey of Service Organizations TOC Literature](#)
[Literature Mapping and Observations](#)
[Limitations of Current Research](#)
[Brief Assessment of Service Management](#)
[What to Change?](#)

[Why Is TOC Not Yet Popular Among Service Organizations' Managers?](#)

[What Do TOC and Focused Management Have To Offer?](#)

[TOC Concepts and Tools for Service Organizations](#)

[The Seven Focusing Steps of TOC](#)

[Bottleneck Management](#)

[Exploiting Permanent Bottlenecks](#)

[Subordinating Everybody Else to the Permanent Bottlenecks](#)

[Elevating the Permanent Bottlenecks](#)

[Response Time Reduction](#)

[Performance Measures](#)

[Costing, Pricing and Decision-Making](#)

[Quality Enhancement](#)

[How to Implement the Change?](#)

[The Remaining Chapters in This Section](#)

[References](#)

[About the Authors](#)

[29 Theory of Constraints in Professional, Scientific, and Technical Services](#) *John Arthur Ricketts*

[Introduction](#)

[Background](#)

[Barriers to Adoption](#)

[Challenges in the PSTS Sector](#)

[What TOC Has to Offer](#)

[What to Change](#)

[Expertise and Assets](#)

[Service Delivery](#)

[Measurement](#)

[Marketing and Sales](#)

[Strategy](#)

[What to Change to](#)

[Replenishment for Services](#)

[Critical Chain for Services](#)
[Drum-Buffer-Rope for Services](#)
[Throughput Accounting for Services](#)
[Nonstandard TOC Applications](#)
[How to Cause the Change](#)
[Buy-in](#)
[How Practitioners Can Get Started](#)
[How Researchers Can Contribute](#)
[What Students Should Know](#)
[Summary](#)
[References](#)
[About the Author](#)

30 Customer Support Services According to TOC

[Alex Klarman and Richard Klapholz](#)

[Introduction—The Need for Change](#)
[What Is Customer Support \(Also Known as Technical Support\)?](#)
[Steady Erosion of Income in the CS Area](#)
[The Warranty Trap](#)
[What to Change](#)
[What to Change to](#)
[A—B](#)
[A—C](#)
[B—D](#)
[C—D'](#)
[D—D'](#)
[Differential Pricing](#)
[The Array of Service Offerings](#)
[Basic Services](#)
[Extended Basic Services](#)
[Limited FSE Visits](#)
[Extended FSE Visits](#)
[Complementing FSE Visits](#)
[Complementing Extended FSE Visits](#)

[Parts Services](#)
[Important Notes](#)
[Other Service Offerings](#)
[Value-Added Services](#)
[Launching of Expert Systems](#)
[Third-Party Maintenance \(or TPM\)](#)
[Installations, Implementations, and Projects](#)
[How to Implement the Change](#)
[Key Decisions](#)
[Policies and Measurements](#)
[Summary](#)
[References](#)
[About the Authors](#)

31 Viable Vision for Health Care Systems Gary Wadhwa

[Introduction](#)
[The Tools for Improvement](#)
[Theory of Constraints](#)
[Lean](#)
[Six Sigma](#)
[Undesirable Effects of the Current Health Care System](#)
[Patients' Perspective](#)
[Doctors' Perspective](#)
[Insurers' Perspective](#)
[Hospitals' Perspective](#)
[Business Owners' Perspective](#)
[Governments' Perspective](#)
[Defining the Goal of the Health Care System](#)
[Improving Quality and Quantity of Patient Flow through Health Systems](#)
[Elaborating on the 5FS](#)
[Thinking Processes for Identifying Root Cause of Physical Constraints to the Flow of Patients](#)

[Throughput Accounting for Performance Measurement and Decision Making in Health Care](#)

[Strategy and Tactic Tree to Implement and Achieve the Viable Vision](#)

[Parallel Assumptions](#)

[Necessary Assumptions](#)

[Sufficiency Assumptions](#)

[An Example](#)

[A Case Study of VV Success](#)

[General Discussion](#)

[References](#)

[About the Author](#)

[Appendix A: Strategy and Tactic Tree for Viable Vision](#)

[Addendum: Excerpt from the Book *Vision for Successful Dental Practice* by Gerry Kendall and Gary Wadhwa](#)

[Steps to success for a private, academic, or government-run dental practice](#)

[32 TOC for Large-Scale Healthcare Systems](#) *Julie Wright*

[Introduction](#)

[Why Change](#)

[Why Healthcare Systems Need to Improve](#)

[The Goal of Healthcare](#)

[What to Change](#)

[Where to Start: Government or Facility?](#)

[The Organic Nature of Healthcare Facilities](#)

[The Human “Engine of Healthcare”](#)

[The Constantly Evolving Workforce](#)

[The Reality of Healthcare](#)

[Current Problem Solving Techniques](#)

[Adapting Industry’s Solutions for Healthcare](#)

[What to Change to](#)

[Where Should the Constraint Reside in Healthcare?](#)

[Starting an Organization on a Process of Ongoing Improvement](#)

[Providing a Safe Platform and an Effective Mechanism](#)

[Building the Current Reality Tree \(CRT\) of a Facility](#)

[How to Cause the Change](#)

[Training the Process Units](#)

[The Process of Ongoing Improvement](#)

[Providing a Knowledge Base for Achieving the Goal Now](#)

[Providing the Knowledge Base for Achieving the Goal in the Future](#)

[Addressing the New Core Problem](#)

[Leaving a TOC Legacy](#)

[Summary](#)

[Proof of Concept](#)

[References](#)

[About the Author](#)

Section VIII TOC in Complex Environments

33 Theory of Constraints in Complex Organizations *James R. Holt and Lynn H. Boyd*

[Overview](#)

[Definition of Complexity](#)

[Major Problems with Complex Organizations](#)

[Undesirable Effects of Complex Organizations](#)

[The Core Conflict for Complex Organizations](#)

[The Direction of the Solution](#)

[What the Market Expects \(A←B\)](#)

[Adding Capabilities \(B←D\)](#)

[Predictable Response to Customers \(A←C\)](#)

[Avoiding Disruptions \(C←D\)](#)

[Doing Both \(D←→D\)](#)

[Additional Understanding of Complex Organizations](#)

[Finding an Injection](#)

[Breakthrough Injection](#)

[Concepts in Organization Complexity](#)

[Categories of Activities](#)

[Flows in Complex Organizations](#)

[Flow Control with Critical Chain](#)

[A Breakthrough Injection](#)

[The Definition of the Common Simple Measure](#)

[Using TDD: An Example](#)

[A Closer Look at the Distribution Department](#)

[Units to Which TDD Applies: Degree of Impact](#)

[on Throughput](#)

[Alternatives for When TDD Does Not Seem to Fit](#)

[Inventory Dollar Days](#)

[Summary of Measures](#)

[Focusing for Balance \(and Changing the Culture](#)

[of the Company.\)](#)

[The Usefulness of Dollar Day Measures in](#)

[General](#)

[A Breakthrough Injection Is Critical, but It Is Rarely](#)

[Sufficient](#)

[Tools for Resolution](#)

[Controlled Resource Allocation](#)

[Challenge of the Future](#)

[The Value of Everyone Measured by the Same](#)

[Simple Measures](#)

[Leadership Certification](#)

[Summary](#)

[References](#)

[About the Authors](#)

[34 Applications of Strategy and Tactics Trees in Organizations](#) *Lisa A. Ferguson, PhD*

[Introduction](#)

[On Becoming an Ever-Flourishing Organization](#)

[The Basic Structure of an S&T Tree](#)

[The Top of the VV S&T Trees](#)

[The Retailer S&T Tree](#)

[Level 2 of the Retailer S&T Tree](#)

[Overview of Level 2 of VV S&T Trees](#)

[Level 3 of the Retailer S&T Tree](#)

[General Overview of the VV S&T Tree Structure](#)

[Levels 4 and 5 of the Retailer S&T Tree](#)

[Need for Lower Levels of an S&T Tree](#)

[Details Regarding the Structure of an S&T Tree](#)

[Key Concepts Regarding Creation of S&T Trees](#)

[How the S&T Tree Relates to Other Thinking](#)

[Process Tools of TOC](#)

[The Other Four Generic VV S&T Trees](#)

[Consumer Goods \(CG\) S&T Tree](#)

[Reliable Rapid Response S&T Tree](#)

[Projects S&T Tree](#)

[Comparison of RRR and Project S&T Trees](#)

[Pay per Click S&T Tree](#)

[Comparison of S&T Tree to Key Literature on Strategy](#)

[Execution of the S&T Tree](#)

[Summary and Discussion](#)

[References](#)

[About the Author](#)

35 Complex Environments *Daniel P. Walsh*

[Introduction](#)

[Brief Background](#)

[Guiding Strategies](#)

[Throughput Accounting](#)

[A Holistic View](#)

[Categories of Variability](#)
[Tools Selection](#)
[A Closer Look at Variability](#)
[Different Tools for Different Types of Variability](#)
[Defining the System](#)
[The TOC Approach](#)
[Applications](#)
[Summary and Discussion](#)
[References](#)
[About the Author](#)

36 Combining Lean, Six Sigma, and the Theory of Constraints to Achieve Breakthrough Performance

AGI-Goldratt Institute

[Introduction](#)
[Lean](#)
[Six Sigma](#)
[Theory of Constraints \(TOC\)](#)
[Discords that can Block the Effective Integration of TOC and Lean Six Sigma \(LSS\)](#)
[Work Behaviors](#)
[Material Release](#)
[Replenishment System](#)
[TOCLSS—Fully Integrated TOC, Lean, and Six Sigma](#)
[References](#)
[About the Author](#)

37 Using TOC in Complex Systems *John Covington*

[Introduction](#)
[We Need More Sucker Rods!](#)
[Introduction](#)
[Some History and What We Learned](#)
[What Change was Needed](#)
[How to Cause the Change](#)

[What We Did to Implement the Change
“Oh Canada”](#)

[Results after Six Months](#)

[Have You Really Defined the System?](#)

[Introduction](#)

[What Do We Need To Change?](#)

[What Do We Change To?](#)

[How Do We Cause the Change?](#)

[Results](#)

[Where is the Constraint in Disciple Making?](#)

[Introduction](#)

[The Analysis](#)

[Results after Two Years](#)

[Summary](#)

[Reference](#)

[About the Author](#)

38 Theory of Constraints for Personal Productivity/Dilemmas *James F. Cox III and John G. Schleier, Jr.*

[Introduction: A Status Report](#)

[Resolving Chronic Conflicts and Developing Win-Win Solutions](#)

[Background: Father-Son Dilemmas](#)

[Personal Productivity Dilemma—Where to Spend Your Time?](#)

[A Review of Constructing the Evaporating Clouds](#)

[College Student Dilemma \(Undergraduate\)](#)

[EC of the Classic Dilemma of White-Collar](#)

[Burnout](#)

[Personal Productivity—Establishing Goals, Strategies, Objectives, Action Plans, and Performance Measures](#)

[What to Change—How Do You Currently Use Your Time?](#)

[Developing a Detailed Implementation Plan to Accomplish Your Goals and Objectives](#)

[Using Buffer Management to Increase Your Effectiveness](#)

[Using the Thought Processes to Achieve Life Goals](#)

[Sheila's Story](#)

[Personal Productivity](#)

[Sheila's Epilogue](#)

[Our Epilogue on Sheila](#)

[Summary](#)

[References](#)

[About the Authors](#)

[Selected Bibliography of Eliyahu M. Goldratt](#)

[James F. Cox III and John G. Schleier, Jr](#)

[Books](#)

[Theory of Constraints Journal Articles](#)

[Journal/Magazine Articles](#)

[Industry Week Late Night Discussion Series](#)

[Management Skills Workshop Series \(Workbooks\)](#)

[Video Movie/Presentations](#)

[Goldratt Program Series \(Video/DVD\)](#)

[Self-Learning Computer Education Software](#)

[Programs](#)

[Necessary and Sufficient Series](#)

[TOC Insights Series. 4 Self-learning Computer](#)

[Software](#)

[Chapters in Books](#)

[Conference](#)

[Proceedings/Video](#)

[Proceedings/Presentations](#)

[Keynote](#)

[Presentations/Video](#)

[Conference](#)

[Presentation](#)

[The Goldratt Webcast Series](#)

[Strategy and Tactic Trees](#)

[POOGI Forum Letter Series](#)
[Plays](#)
[Commercial Software](#)

[**Index**](#)

Preface

Beginning in the early 1980s with the OPT software, a software package for scheduling manufacturing operations, Dr. Eliyahu M. Goldratt started applying the concepts of the hard sciences¹ to problems in organizations. Later, with the publication of *The Goal* in 1984, Dr. Goldratt launched a series of revolutionary concepts aimed at bringing about improvement in the global performance of organizations by *focusing on a few leverage points* of the system. These revolutionary ideas of Theory of Constraints go to the very core of how things work in the real world. They focus on *Constraints* as a centerpiece in the definition and management of production work flow in manufacturing, administrative processes, project management and the like. *Holistic thinking* is emphasized throughout, shifting the focus on work direction and measurement from local efficiencies to *Throughput* of the entire system, *buffering* the system to protect it from the statistical fluctuation caused by unexpected problems (Murphy), Parkinson's Law, etc. This is fortified with clear guidance on placement of buffers in the flow of the system and simple tools for "*Buffer Management*" as a way of achieving the best focus on priority actions. By taking a systems view and focusing the cause-and-effect relationship of the leverage points to global performance, Goldratt invented new management concepts and applications in production, project management, finance, accounting and performance measurement, distribution and supply chain, marketing, sales, managing people, and strategy and tactics. The concepts are robust with applications appearing in manufacturing, services, engineering, government,

education, medicine, prisons, banking, and professional, scientific, and technical services and other service industries.

Perhaps Dr. Goldratt's most important contributions are the Thinking Processes which employ structure and language to lay out true cause and effect in defining problems and laying out conflict dilemmas and their solutions. They have been taught and used effectively in all levels of education from pre-kindergarten through PhD research. On a grand scale they provide a suite of complementary problem-solving and decision-making tools based on using the scientific foundation of cause-and-effect logic, with steps for verification and validation. While they are applied in strategy, development, marketing, sales, production, distribution, finance, and accounting, they are useful for addressing personal problems and have even been used in teaching prisoners how to deal with the issues they face.

Theory of Constraints concepts and tools are aimed at one overriding objective: bringing about a process of ongoing improvement in enterprises. That said, the purpose of this book is to provide "hands on" guidance from the world's top experts on how to implement these TOC capabilities. This guidance is buttressed by clear definition on how they work, why they work, what issues are resolved and what benefits accrue. Leading practitioners provide guidance based on their hands-on implementation experience. Academic authors give a review of the wealth of literature on why to move from the traditional discipline to each TOC discipline and a review of the TOC literature in that discipline. Indeed these ideas are of such a scope that this Handbook required 44 authors to explain them.

*James F. Cox III
John G. Schleier, Jr.*

Acknowledgments

We have a number of people and organizations to acknowledge for their help in bringing this handbook to the public. First and foremost are our wives, Mary Ann and Maribeth; they deserve all of the credit as they managed our daily lives so we could work more than full time on this project. They lived through our ups and downs with us at every step of this journey. Second, we appreciate our children and grandchildren for understanding that we had a major project at hand and hopefully fun would come later.

Dr. Goldratt deserves a lot of praise and appreciation for breaking the ground on topics of this handbook in his drive to teach the world to think logically. For thirty years he has been slaying sacred cows in the various business functions. From the very beginning he had a focus on the goal of teaching people to think like scientists and on redefining the system perspective to mean identifying and managing a few control points in a system and where change makes a significant difference in system performance. From our perspective, his biggest contributions are his Thinking Processes. They offer the potential of teaching children and adults right from wrong, how to sort through their personal and business problems, and how to achieve their dreams by using simple logic. He emphasized the use of the Socratic approach as a powerful methodology for both teaching and gaining buy-in.

The TOC experts (and their supportive families and friends) who authored these chapters must be thanked and applauded. They worked diligently at writing so we, the editors, could understand their shared wisdom. Their depth-of-knowledge is unparalleled. Each worked long hours

through many revisions of their chapter to achieve a lasting contribution to the body of knowledge. But let us all understand that tomorrow there will be more and better material in each area as we have embarked on a journey and have not arrived at a destination with this handbook. It is a stake in the ground: what we know now. Let's move forward.

Some professional organizations and individuals should be singled out. The Theory of Constraints International Certification Organization (TOCICO) is the young but rapidly growing certifying body of TOC, holding regional and international conferences focusing on the development of TOC knowledge and certification of professionals in this knowledge. Their *TOCICO Dictionary* has been a valuable source for definitions of key terms. Second, APICS is one of the leading educational and certification organizations, and the first to explore, and present education in TOC. By providing access to their excellent dictionary for a modest fee, APICS has enabled us to enhance the quality of the handbook. The Ansoff Family Trust permitted the use of the Ansoff's growth matrix. TOC for Education (TOCfE) deserves special recognition for its efforts worldwide to teach children to apply logic and common sense in their life. Kathy Suerken, its director, has been a torch bearer for its cause and these professionals. We also thank the Singapore Prison system for permission to publish their unique and positive experience with TOC. We also thank Eli Goldratt, North River Press, and John Thompson for use of their materials. Wendy Donnelly and Jennifer Tucker provided valuable assistance in documenting references and building the bibliography.

Our education in TOC goes back almost to the very beginning with Dr. Goldratt and Bob Fox of Creative Output speaking at early APICS Conferences and offering academics invitations to attend their workshops. The list of TOC professionals extends through the Avraham Y. Goldratt Institute with Bob Fox, Dale Houle, Eli Schragenheim, Shri

Srikanth, Oded Cohen, Alex Klarman, Alex Mashar, Tracey Burton Houle, Dee Jacobs, Debra Smith, John Covington, and many others. Many of the authors in this handbook are past students of Eli Goldratt and AGI and have since become leaders in our field. These individuals have freely shared their knowledge with us over the years. We have asked these individuals to share their knowledge now with you. We honor them for doing such an extraordinary job.

Theory of Constraints Handbook

SECTION I

What Is TOC?

CHAPTER 1

Introduction to TOC—My Perspective

Here Dr. Goldratt, developer of TOC gives his perspective on what TOC is; its goals and objectives, and the state of its progress in bringing about improvement. Dr. Goldratt discusses the evolution of TOC: how the identification of major system problems led to the development of solutions and significant system improvement only to surface the next system problem. . . . Thus the evolution of TOC followed the natural scientific approach to system improvement. As the developer of Theory of Constraints, he has brought the mind of a scientist to the problems and needs of business, private sector organizations, and individuals. His scientific approach has led to the breaking of several business paradigms and the development of new simplified approaches to managing systems. In the section, his chapter leads forward to the remainder of the book where the depth and scope of the TOC concepts are seen in action.

CHAPTER 1

Introduction to TOC—My Perspective

Eliyahu M. Goldratt

There is a famous story about a gentile who approached the two great Rabbis of the time and asked each, “Can you teach me all of Judaism in the time I can stand on one leg?”

The first Rabbi chased him out of the house, however, the second Rabbi answered: “Don’t do unto others what you don’t want done to you. That is all of Judaism, the rest is just derivatives. Go and learn.”

Can we do the same; can we condense all of TOC into one sentence? I think that it is possible to condense it to a single word—focus.

Focus

There are many different definitions to the word *focus*, but a good starting point is a simple definition such as “Focus: doing what should be done.”

Focusing on everything is synonymous with not focusing on anything.

In almost any system, there are plenty of actions that will contribute to the performance of the system, so what is the difficulty in focusing? True, we can't take all the beneficial actions because we don't have enough time or enough money or enough resources, but the more we do, the better it is. This naïve view was shattered by Pareto¹ with his 80-20 rule. What Pareto proved is that 20 percent of the elements contribute 80 percent of the impact. Therefore, when we can't do it all, it is of the utmost importance to properly select what to do; it is of the utmost importance what we choose to focus on.

Copyright © 2010 by Eliyahu M. Goldratt.

However, as Pareto himself pointed out, the 80-20 rule is correct only when there are no interdependencies between the elements of the system. The more interdependencies (and the bigger the variability), the more extreme the situation becomes. In organizations, there are numerous interdependencies and relatively high variability; therefore, the number of elements that dictate the performance of the system—the number of constraints—is extremely small. Using Pareto's vocabulary, one might say that in organizations 0.1 percent of the elements dictate 99.9 percent of the result. This realization gives new meaning to the word *focus*.

Constraints and Non-Constraints

There isn't a more grave mistake than to equate non-constraint with non-important. On the contrary, due to the dependencies, ignoring a non-constraint can impact the constraint to the extent that the performance of the entire system severely deteriorates. What is important to notice is that the prevailing notion that "more is better" is correct

only for the constraints, but it is not correct for the vast majority of the system elements—the non-constraints. For the non-constraints, “more is better” is correct only up to a threshold, but above this threshold, more is worse. This threshold is dictated by the interdependencies with the constraints and therefore cannot be determined by examining the non-constraint in isolation. For the non-constraints, local optimum is not equal to the global optima; more on the non-constraints does not necessarily translate to better performance of the system.

An hour lost on the bottleneck is an hour lost on the entire system; an hour gained on a non-bottleneck is a mirage.

We now recognize that the vast majority of the elements of a system are non-constraints. We also recognize that for non-constraints more might not be better but worse. So, what must be the unavoidable result of following the prevailing notion that more is better? The number one reason for not doing what should be done is doing what should not be done.

We don't have a choice but to define *focus* more narrowly: do what should be done AND don't do what should *not* be done.

Measurements

According to cost accounting, when operations produce they absorb cost into the inventory and this cost absorption is interpreted as increasing profit. In other words, the cost-accounting concept encourages any production, even on a non-bottleneck, even if it is above the threshold. It is no wonder then that the first implementations of TOC clashed with cost accounting. It was mandatory to develop an alternative. Almost immediately, Throughput Accounting

(TA)—a system based on simple definitions of Throughput (T), Inventory (I), and Operating Expense (OE)—was proposed alongside the explanation of the difference² between the Cost World and the Throughput World.

Tell me how you measure me and I will tell you how I will behave.

The Goal and The Race

Rapidly, the realization of the crucial impact of bottlenecks gave rise to a collection of actions that were previously deemed inefficient and now were recognized as the most important actions to be taken. “What should be done” now took on new meaning.

A novel about manufacturing? We don’t even know what shelf to put it on. It will never work.

No less important was the recognition that it is impractical to monitor each non-bottleneck separately and therefore constructing and implementing a system to prevent the overproduction of non-bottlenecks was essential (Drum-Buffer-Rope [DBR] and Buffer Management [BM]). The understanding of “What should not be done” was even more tantalizing.

This body of knowledge was captured in detail in *The Goal* (Goldratt and Cox, 1984) and conceptually explained in *The Race* (Goldratt and Fox, 1986).

Other Environments

The clear logic, the simplicity, and the rapid results that TOC provided in production caused other environments to try to implement the same. Unfortunately, some of them were so different that even the constraint was different in nature. The constraint in project environments is not bottlenecks but the critical path (or, more accurately, the critical chain). The constraint in distribution has nothing to do with bottlenecks. It is either cash (wholesalers) or the number of clients that enter the shop (retail). The term *bottleneck* started to be misleading; it had to be replaced with the broader term *constraint*. That was the time (1987) when the term *Theory of Constraints*³ was coined and a precise verbalization of the focusing process was offered—the five focusing steps.

When you are good with a hammer everything looks like a nail.

That was not enough. Applications for proper guidance of the non-constraints in distribution⁴ (blocking the tendency to push the merchandise downstream [replenishment to daily consumption]) and in project environments (blocking the tendency to buffer the individual tasks [Critical Chain Project Management]⁵) had to be developed in full.

The Thinking Processes⁶

Only when environments other than production had been dealt with using TOC did the paradigm shift dictated by the narrower definition of focus fully surface. To focus properly, the following questions had to be answered: How do we identify the constraint? What are the decisions that will lead to better exploitation? How do we determine the proper way to subordinate the non-constraints to the above decision?

And how do we reveal more effective ways to elevate the constraint? It became apparent that even the best available practices were not delivering the required answers, and relying on intuition was not enough.

Reality is exceedingly simple and harmonious with itself.

The standard ways to identify the needed actions, the standard ways to focus the improvements, were obviously not adequate. They usually started with a list of problems, of gaps between the existing situation and the desired situation. The gaps were quantified and, following the Pareto principle, items at the top of the list were taken as the targets for improvement.

This approach leads, at best, to just marginal improvements, since at the base of the approach is the erroneous assumption that the gaps are not interdependent. When the interdependencies are taken into account, it becomes apparent that the gaps are nothing but effects, *undesirable effects* (UDEs) of a much deeper cause. Trying to deal directly with the UDEs does not lead to the recognition of what actions should be taken. Actually, it leads to many actions that should *not* be taken. There was a crying need to provide a logical, detailed structure to identify the core problem, to zoom in on the ways to remove it, and to do so without creating new UDEs. From 1989 to 1992, the Thinking Processes of TOC were successfully developed and polished.

The Market Constraint

When TOC is implemented in operations, the improvements are substantial to the extent that the constraint moves into the market. Very early on, it was noticed that the improved

performance of operations opened new opportunities to gain more sales. That situation was described in *The Goal* (Goldratt and Cox, 1984). But it took several years, and many successful implementations, until it dawned on me that the improvements in operations not only open new opportunities but actually provide the company with a decisive competitive edge. When the constraint of a company is in the market and at the same time the company has a decisive competitive edge, the obvious interpretation of *focus* is to concentrate on capitalizing on the existing competitive edge, rather than being distracted with ongoing refinement in operations. To provide the bridge from the focus on operations to the required focus on strategy, *The Goal* (1992) was extended.

A decisive competitive edge is gained only when a company satisfies a significant market need to an extent that none of its significant competitors can.

To gain the required focus, a clear verbalization of the resulting competitive edge was needed. That wasn't a triviality. What obscured the picture was the fact that the same improvements in operations gave rise to not one but many vastly different competitive edges (depending on the company's products and the nature of their clients). In *It's Not Luck* (Goldratt, 1994), some examples of competitive edges were described, alongside the introduction of the Thinking Processes.

Capitalize and Sustain

Surprisingly, most companies that implemented TOC in operations did not move on to capitalize on the resulting competitive edge. In other words, they became totally unfocused, being satisfied with the results of improved

operations and blind to the much bigger gains that were now readily available—the profit increase when much more sales are won and serviced with already exposed excess capacity. What was missing was a whole body of knowledge.

Be careful of what you wish for. (You might get it. Too much, too soon.)

Rarely does a company have a decisive competitive edge. No wonder that most sales people are not trained to conduct a sales meeting when they do have a decisive competitive edge. The nature of such sales meetings is different from conventional meetings. Rather than concentrating on the company's products, the meeting should revolve around the client environment, exposing its significant need that currently isn't satisfied by the vendors. Since there are many clients' environments, deciphering the causes and effects that govern each of them, constructing the sale cycle in accordance, and finding the way to take the sales people through the required paradigm shift took several years.

But with the first successful cases it became apparent that we had to deal with another challenge. Capitalizing effectively on a decisive competitive edge causes the sales to increase sharply. The resulting jump in sales can easily cause the constraint to swing back into operations—bottlenecks swiftly reappear. If this bounce-back is not properly controlled, it can demolish the competitive edge. To continue to focus, it became essential to know how to sustain the increase in sales and how to synchronize between sales and operations so that the rate of incoming orders will not collapse but continue to grow. It wasn't difficult to figure out the simple mechanisms that provide such synchronization, but it was difficult to face the fact that to actually implement them the TOC implementation had to be done holistically. At that stage, I underestimated the

difficulty of moving from a functional implementation into a holistic implementation and I naively assumed that showing that TOC covers all aspects of the organization would be sufficient. The Satellite Program (Goldratt, 1999), the summary of the TOC knowledge in eight sessions⁷ of 3 hours each, was recorded with that purpose in mind.

Ever Flourishing

A Process of Ongoing Improvement (POOGI) was the subtitle of the revised edition of *The Goal* (Goldratt and Cox, 1986) and the motto of TOC. Early on, it was noticed that the conventional definition of POOGI (performance goes up as time advances) contains two conceptually different curves⁸—the red curve, where the rate of improvement grows leading to exponential growth, and the green curve, where the rate of improvement decays leading to diminishing returns. The drive to move companies to capitalize on the resulting competitive edge that stems from the improvement in operations caused us to guide companies to strive for the red curve and to condemn the green curve.

The biggest obstacle to achievements—setting the objective too low.

Only when reality demonstrated the absolute necessity of sustaining rapid growth did it dawn on me that the green curve is as essential as the red curve. Actually, we are dealing with two types of performances: financial growth and stability. Companies should strive to ensure that their financial performance will grow by at least a few percent per year, which is equivalent to demanding the red-curve growth. But, to ensure that such growth will be sustained, companies must ensure that the growth will not degrade their stability. It became more and more evident that

achieving the red curve mandates the attainment of the green curve and vice versa.

To “make more money now as well as in the future” (the objective stated in *The Goal*) it is essential to choose carefully the actions that will not only bring growth in the near future but also increase (rather than endanger) the company’s stability for the longer horizon. To fully capture this essential realization, the objective was rephrased to: “become an ever-flourishing company.” Likewise, the paths to reach an ever-flourishing stage had to be laid out in detail. Focus, doing what should be done and not doing what shouldn’t be done, forced us to, again, re-examine and severely alter the conventional wisdom.

At that stage (2002), the knowledge already existed, in sufficient details, to construct the paths for five different types of industries: make-to-order, make-to-stock, project based, equipment manufacturers, and retailers/wholesalers. This knowledge was so vast that it took many years to educate new experts. Even more troublesome was the fact that the transfer of even the relevant section of the knowledge needed for improving a specific company raised numerous misunderstandings. A comprehensive tool for clearly transferring a vast body of knowledge was mandatory.

Strategy and Tactic Trees

The *Strategy and Tactic tree* (S&T) is probably the most powerful tool of the Thinking Processes. Formally, it replaces the prerequisite tree. Practically, it is the organizer of all the knowledge gained by the previous tools. It is the logical structure that enables focusing. Starting from the company’s strategic objective, it logically derives what

actions (and in which sequence) must be taken and which actions should not be taken.

Strategy—the answer to “What for?” Tactic—the answer to “How?”

The S&T trees brought clarity to the implementations. They enhanced communication through the management levels and synchronization between the various departments. The time to reach results was considerably shortened and the transition, from one stage of implementation to the next, became relatively smooth. No less important, they enabled introducing this knowledge (the detailed implementation plan for the five environments)⁹ into the public domain. That was accomplished through a series of (recorded) Web seminars in 2008–2009 (Goldratt, 2008; 2009)¹⁰.

New Frontiers

Currently several important new frontiers are screaming for answers. And I suspect that this will always be the case as long as we continue to be good scientists. My opinion about it has not changed in the last 25 years. So, maybe the best way to summarize this introduction is by quoting, word by word, from my introduction to *The Goal*:

The secret of being a good scientist, I believe, lies not in our brain power. We have enough. We simply need to look at reality and think logically and precisely about what we see. The key ingredient is to have the courage to face inconsistencies between what we see and deduce and the way things are done. This challenging of basic assumptions is essential to breakthroughs. Almost everyone who has worked in a

plant is at least uneasy about the use of cost accounting efficiencies to control our actions. Yet few have challenged this sacred cow directly. Progress in understanding requires that we challenge basic assumptions about how the world is and why it is that way. If we can better understand our world and the principles that govern it, I suspect all our lives will be better.

A powerful answer raises new fruitful questions.

References

- Blackstone, J. H. 2008. *APICS Dictionary*. 12th ed. Alexandria, VA: APICS.
- Goldratt, E. M. 1990. "[Chapter 6](#): The paradigm shift." *The Theory of Constraints Journal* 1(6): 1-23.
- Goldratt, E. M. 1994. *It's Not Luck*. Great Barrington, MA: North River Press.
- Goldratt, E. M. 1997. *Critical Chain*. Great Barrington, MA: North River Press.
- Goldratt, E. M. 1999. *Goldratt Satellite Program*. Video Sessions 1-8. Brummen, The Netherlands: Goldratt Satellite Program.
- Goldratt, E. M., Schragenheim, E. and Ptak, C. A. 2000. *Necessary but not Sufficient*. Great Barrington, MA: North River Press.

- Goldratt, E. M. 2008. *The Goldratt Webcast Series: Critical Chain Project Management*. Roelofarendsveen, The Netherlands: Goldratt Marketing Group.
- Goldratt, E. M. 2009. *The Goldratt Webcast Series: From Make-to-Stock (MTS) to Make-to-Availability (MTA)*. Roelofarendsveen, The Netherlands: Goldratt Marketing Group.
- Goldratt, E. M. and Cox, J. 1984. *The Goal: Excellence in Manufacturing*. Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. and Cox, J. 1986. *The Goal: A Process of Ongoing Improvement*. rev. ed. Crotonon-Hudson, NY: North River Press.
- Goldratt, E. M. and Cox, J. 1992. *The Goal: A Process of Ongoing Improvement*. 2nd rev. ed. Great Barrington, MA: North River Press.
- Goldratt, E. M. and Fox, R. E. 1986. *The Race*. Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. and Fox, R. E. 1987. *The Theory of Constraints Journal*, volume 1.
- Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>

About the Author

Eli Goldratt is known by millions of readers worldwide as a scientist, educator, and business guru. His Theory of Constraints (TOC) is taught at business schools and MBA programs around the globe. Government agencies and

businesses, large and small, have adopted his methodologies. TOC has been successfully applied in almost every area of human endeavor, from industry to healthcare to education. And while Eli Goldratt is indeed a scientist, an educator, and a business leader, he is first and foremost a philosopher; some say a genius. He is a thinker who provokes others to do the same. Often characterized as unconventional and always stimulating—a slayer of sacred cows—Dr. Goldratt exhorts his readers to examine and reassess their lives and business practices by cultivating a different perspective and a clear new vision.

SECTION II

Critical Chain Project Management

CHAPTER 2

[The Problems with Project Management](#)

CHAPTER 3

[A Critical Chain Project Management Primer](#)

CHAPTER 4

[Getting Durable Results with Critical Chain—A Field Report](#)

CHAPTER 5

[Making Change Stick](#)

CHAPTER 6

[Project Management in a Lean World—Translating Lean Six Sigma \(LSS\) into the Project Environment](#)

Projects are at the center of change in organizations. They are the vehicles for new product development, major process improvements, organization changes, and the like. Organization strategies therefore depend on projects for their execution making it vital that projects be carried out in the most effective way possible.

As the chapters of this section reveal, Theory of Constraints Critical Chain unlocks a series of new paradigms that enable major advances over traditional methods. New approaches consider availability of critical resources in the

timing of new project releases and the planning of individual project schedules. New concepts in task estimating and tracking open the door to intelligently placed protective time buffers enabling managers to focus correctly on specific areas that need attention for project success. Elimination of unnecessary multi-tasking combines with a “relay-runner” approach to work flow to dramatically reduce project execution times and improve project quality. These simple but effective concepts focus management and resource efforts on the vital few tasks that determine organization success.

Key steps to implementation and sustainability are addressed. These techniques and the dramatic improvements in the field are explained. These include huge improvements in completing projects on time, to specification, and within budget. This section gives a clear picture of the Critical Chain concept and how to implement and execute it. Integration of Critical Chain with Lean and Six Sigma is included. While management of individual projects is addressed, special emphasis is put on multi-project environments as these are more pervasive.

CHAPTER 2

The Problems with Project Management

Ed Walker

Introduction

Most projects fail! Failure generally means the actual results of at least one of the three objectives of the project did not meet original expectations. The project scope was reduced (changed the original specifications), the project was delivered late (compared to the original due date), or the budget was exceeded (actual project costs exceeded projected costs). In some projects, two or even all three of these objectives were not realized.

Over the last four decades, two streams of research have emerged from project management. In the management science stream, numerous academic researchers have studied project networks (the theory) to identify specific problems with PERT/CPM (use of the Beta distribution, limited resources, parallel paths, etc.) or to determine the most efficient algorithm to identify the shortest time to complete a project. In the management arena, numerous academic researchers and practitioners have studied the project management environment to identify human problems (lack of project and technical skills, lack of teamwork, lack of communications, etc.) as causes of project failure. Seldom have these researchers acknowledged the work in the other research stream as

possible causes for project failure. In many cases, the management scientists only discussed the problem under study as the cause of project failures. What is required, then, is an examination of the project environment as a system and determination of the causes of project failure.

Purpose and Organization of the Chapter

The overarching purpose of this chapter is to expose the problems associated with “traditional” project management. Specific solutions are not to be found here, but rather what is provided is a framework for developing a new project management method that uses a systems perspective to address the core problems of traditional project planning, scheduling, and controlling tools. A systems perspective is needed to fully assess the impact of an assumption related to an activity, to resource contention, or to converging paths on the results of a project.

Copyright © 2010 by Ed Walker

To accomplish this task, first the reader will be provided with brief overviews of both Gantt chart scheduling and PERT/CPM. Gantt charts were first developed nearly 100 years ago, while PERT (initially named *Program Evaluation Research Task* and later changed to *Program Evaluation and Review Technique*) and CPM (*Critical Path Method*) began their evolution approximately 50 years ago. Each of these methods has both advantages and disadvantages, which are summarized. Brief reviews of the literature related to the origins of project management and the high failures of projects since its origin, as well as literature related to both single and multiple project networks and resource allocation are then presented. The main body of the chapter is devoted to the development of guidelines that any new

project management method must address. This is followed by a brief introduction to *Critical Chain Project Management* (CCPM) and finally a review of the recent CCPM literature.

Traditional Planning and Control Mechanisms in Project Management

Gantt Charts

A Gantt chart is a horizontal bar chart developed as a production control tool in 1917 by Henry L. Gantt, an American engineer and social scientist. Frequently used in project management, a Gantt chart provides a graphical illustration of a schedule that helps to plan, coordinate, and track specific activities in a project. These charts might be as simple as a hand-drawn image on graph paper, or as complex as purpose-built computer software. A simple Gantt chart used for a project is shown in [Fig. 2-1](#).

The horizontal axis of the chart represents the total time span of the project (broken down into uniform time increments—days, weeks, months, etc.), while the tasks comprising the project are on the vertical axis. Horizontal bars are used to illustrate the start and end dates of individual activities (for example, task A has a duration of y days starting on day 1 and ending on day 5). In its simplest form, the Gantt chart shows all of the activities necessary to complete the project. Some of the activities must be completed in a specified sequence, while others might progress concurrently. Tasks B and C are processed sequentially and tasks B and D can be processed concurrently. One cannot start framing a home before the foundation is laid; but once framing is complete, the

plumbing and electrical systems can be installed simultaneously.

More complex Gantt chart scheduling is often based on a *work breakdown structure* (WBS). To continue the previous example, the installation of the electrical system (the objective) might be broken down into manageable elements such as the installation of the breaker panel, pulling electrical wires through the home, pulling data cables, connecting the electrical wires to the breaker panel, inspection of the system by the building inspector, etc. It is then necessary to determine the start and end dates as well as responsibility for each. In such a chart, percent completion is tracked for each element and the objectives. A vertical line on the chart shows the current date (March 25) while the completed and noncompleted portions of each horizontal bar are shaded differently to allow visual inspection of the project's progress. For example, task B is late by two days and task D is early by one day.

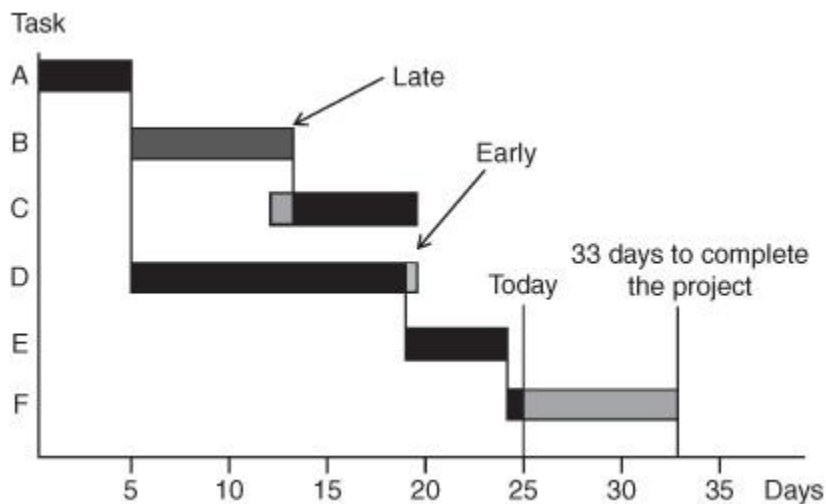


FIGURE 2-1 Simplified Gantt chart of a project network.

The primary advantages of Gantt chart scheduling are that it can be easily understood by a wide audience and it provides a visual means to track project progress. The

disadvantages are numerous. The chart becomes unwieldy for larger projects (more than about 30 activities) when it extends for more than a single page (or screen, if computerized). The chart does not indicate task dependencies, and therefore fails to communicate how falling behind on one activity might affect other activities. When using a WBS, often there is confusion between defining the WBS and defining the activities of the project. Additionally, as some elements of the WBS might be front- or end-loaded (more work at the beginning or end of the element), the percent progress reported might be over- or understated.

PERT/CPM in the Single Project Environment

CPM and PERT originated in 1957 and 1958, respectively, with CPM examining the tradeoffs between project duration reduction and increases in activity and project costs; and with PERT examining the uncertainty aspects of completion dates for development projects. CPM was originally developed for use with manufacturing plant rebuilds by DuPont and PERT for use with the Polaris nuclear submarine program by the Special Project Office of the Department of the Navy and the consulting firm Booz Allen Hamilton. From their origins to the present, both techniques (and their subsequent merger into one) have been heralded as breakthroughs in managing complex systems.

Once all of the activities are identified (a process that in itself is subject to controversy), a project network can be created. The network organizes the activities in such a way as to clearly show the technological precedence relationships—the simple fact that most activities must be preceded or followed by some other activity or activities. [Figure 2-2](#) shows a typical activity-on-arrow project network. This network has six activities, each having an associated

estimate of activity duration. PERT/CPM requires that a forward pass through the network be made to determine the *early start* (ES) and *early finish* (EF) times for each activity. Then a backward pass is made through the network using the EF time of the last activity as the *late finish* (LF) time of the last activity. This backward pass determines the late finish and *late start* (LS) of each activity. The difference between the LF and EF (or LS and ES) is the slack associated with each activity.

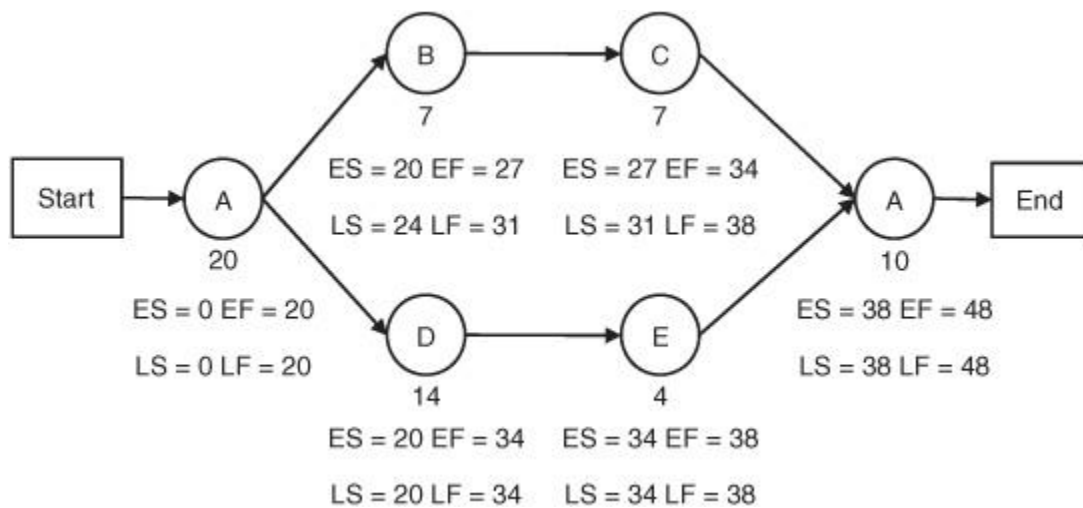


FIGURE 2-2 Typical activity-on-node project network.

Activities having zero slack are called critical activities because any delay in these activities will cause the project to be late. In [Fig. 2-1](#), the critical activities are A, D, E, and F. These activities form what is referred to as the critical path. The primary advantages of PERT/CPM over Gantt charts are that technological precedence (activity dependency) is readily apparent and that it is relatively easy to determine how falling behind on one activity might affect other activities. Activities B and C ([Fig. 2-2](#)) have a calculated slack value of four. Therefore, if either of these activities is delayed by more than four days, then the critical path will be negatively impacted because both activities C and E

must be completed before activity F can be started. The primary disadvantage of this method is the assumption of readily available capacity on the required resources. Much of the research that follows expands upon this basic premise.

Brief Review of Project Management Literature

The project management literature is enormous—several thousand articles and dozens of books. While project management has grown significantly, many of the problems initially identified almost five decades ago in both the macro-view (applications) literature and in the micro-view (or theoretical) literature still exist today. This literature review provides only a brief glimpse of the continuing themes of project management research since the late 1950s to today. The purpose is to show that the problems of project management have not been solved nor has the promise of project management been achieved over this time interval.

Origins of PERT and CPM

The project management literature is intertwined with descriptions of the benefits and problems of PERT and CPM. A brief account is provided illustrating this continuing dialog. PERT and CPM successes immediately caused top management interest. Several researchers provided articles describing the use of these new management tools. Malcolm, Rose-boom, and Clark (1959) provide a status report with a history of the development; examples of the flow plan; the elapsed-time estimates; the organization of the data; the computation of expected times, latest times, slack, and critical path; and probability of project completion

by a given date. A description of the pilot application and its full-scale implementation and results to date are provided as well. Almost as soon as this detailed account of PERT appeared, Healy (1961) warned of a problem with the technique, that subdividing activities and their related times can change the project completion date probabilities. Clark (1961) and Millstein (1967) critique Healy's research based on the realities of managing with PERT, while Roseboom (1961) critiques Healy's research based on the whether his assumptions are realistic.

Miller (1962) provides a description of how to plan and control with PERT; and Levy, Thompson, and Wiest (1962) provide a similar description of the ABCs of CPM, both in the *Harvard Business Review*. At the same time, Pocock (1962) describes PERT, its payoff, and problems (PERT is a management responsibility; PERT is no automatic system; PERT often clashes with traditional organization patterns; learning to use a dynamic control system; poor applications; PERT cannot be a rigidly standardized technique). Kelley (1962) provides research supporting the mathematical basis of CPM, and Bildson and Gillespie (1962) provide an extension with a model of PERT (with activity time uncertainty) and CPM (with cost of activity crashing). Paige (1963) provides a detailed description of PERT/Cost. Several research articles followed which examined and attacked the PERT assumptions as being unrealistic or false. The purpose of providing a brief description of the early articles and the articles in the appendices related to various micro (theory) problems being argued even today is to illustrate the types of causes that academic researchers identify for the failures of projects.

Project Failures

In addition to the theory-based research, both anecdotal articles and surveys of different types of project organizations have been taken to determine the level of project failures and causes of failures. In 1957, C. Northcote Parkinson observed that “work expands so as to fill the time available for its completion”—now known as Parkinson’s Law. Others (Marks and Taylor, 1966; Krakowski, 1974; Gutierrez and Kouvelis, 1991) have ascribed the presence of this law to project activities and the results on project duration.

Middleton (1967) surveyed project management organizations in aerospace industries. Respondents provided the following disadvantages of using a project management organization: more complex internal operations (51 percent), inconsistency in application of company policy (32 percent), lower utilization of personnel (13 percent), higher program costs (13 percent), more difficult to manage (13 percent), and lower profit margins (2 percent). Other disadvantages provided were the tendency of functional groups to neglect their jobs, too much shifting of personnel from project to project due to priorities, and duplication of functional skills in the project organization.

Avot states, “The many instances where project management fails overshadow the stories of successful projects” (1970, 36). He identified the major causes of project failure as the following: the basis for the project is not sound; the wrong man is chosen as project manager; company management is unsupportive; tasks are inadequately defined; the project management system is not adequately controlled; management techniques (e.g., too many reports) are misused; and project termination is not planned.

Brooks (1995), the project manager for the IBM OS/360, provides five major causes for lateness in information technology projects: (1) techniques of estimating are poorly developed (estimates are usually optimistic); (2) estimating

techniques confuse effort with progress (the assumption is that men and months are interchangeable); (3) submission to the customer's desired (but unrealistic) due date; (4) schedule progress is poorly monitored; and (5) when schedule slippage occurs, the response is to add manpower.

Based on his project management experience, Hughes (1986) blames the majority of project failures on not following basic management principles such as an improper focus on the project management system instead of the project goals; fixation on maintaining first-time estimates; too detailed or too broad activity structure; lack of training in project management techniques; too many people assigned to the project (Parkinson's Law); lack of communication of goals; and rewarding the wrong actions. Black (1996) surveyed professional engineers to determine the causes of project failures. His top 12 causes of project failures are:

1. Lack of planning
2. The project manager
3. Project changes (scope creep, poor planning, etc.)
4. Poor scheduling
5. Skills of team members
6. Management support
7. Funding
8. Cost containment
9. Resources
10. Information management

11. Incentives (lack of rewards and penalties)

12. Lack of continuing risk analysis

In the late 1980s and early 1990s, a series of articles by Pinto and Slevin (1987), Pinto and Prescott (1988), Pinto and Slevin (1989), and Pinto and Mantel (1990) examined the presence of critical success factors in project implementations; differences across the stages of the project life cycle; differences between construction and R&D projects; and project failures, respectively. The critical success factors are project mission definition, top management support, client consultation, personnel, technical tasks, client acceptance, monitoring and feedback, communication, and troubleshooting. Of the projects (Pinto and Mantel, 1990) assessed as failures in the strategic stages, the relevant criterion for failure was that related to external effectiveness: perceived value of the project and client satisfaction. In the tactical stages, the relevant criteria for failure were those related to trouble shooting, lack of adequate personnel, ineffective scheduling, lack of client acceptance, and inadequate technical support. One hypothesis (H3), "the perceived causes of project failure will vary depending upon the type of project assessed: construction or R&D" (1990, 271), proved to be true. In construction projects, the causes of failure were lack of technical expertise, support, and lack of adequate trouble shooting mechanisms. In R&D projects, a wide variety of causes of failure was identified with the cause depending on the definition of failure; inadequate troubleshooting impacts all definitions; implementation process: inefficient scheduling; client satisfaction: personnel and monitoring and feedback; and quality: lack of clear statement of project goals.

Brown (2001) reports that three-quarters of all projects are completed late and over budget according to a survey of 1800 executives, practitioners, and consultants. Pitagorsky

(2001) puts the failure rate at 40 to 60 percent. According to James, 40 percent of all large IT projects end in “utter failure” while another 33 percent are “challenged, meaning that they were completed late, over budget or with fewer features and functions than originally specified” (2000, p. 40). Based on their 20 years of project management experience, Matta and Ashkenas (2003) provide two major causes of complex project failure—critical tasks (called white-space risk) are left off the project plan and the different activities won’t come together in the end to produce the final project.

Neimat (2005) provides a detailed analysis of IT project failure research from the Standish Group by annual IT surveys of 18,000 executives showing the trends in failure rates from 1994 (over 80 percent of projects were challenged or failed) through 2000 (about 70 percent of projects were stressed or failed); a summary of several more recent research efforts examining IT project failures and a description of the Federal Bureau of Investigation (FBI) Virtual Case File project failure. His listing of causes of project failure is similar to the listings across the 40-year period: poor planning, unclear goals and objectives, objectives changing during the project (scope creep), unrealistic time or resource estimates, lack of executive support and user involvement, failure to communicate and act as a team, and inappropriate skills. Interestingly, the descriptions of the failure and success variables listed in the various articles across this 40-year time period are quite similar.

Single Project Management Literature

PERT/CPM is criticized for failing to provide achievable completion dates, for consistently underestimating budgets, and for using resources inefficiently (e.g., Klingel, 1966;

Badiru, 1993; Meredith and Mantel, 2003, 134-135, 649-652). These failures might be traceable to a faulty initial plan or to an inadequate control process. A variety of methods for both planning and controlling projects has been espoused by researchers (Wiest and Levy, 1977; Badiru, 1993; Kerzner, 1994; Meredith and Mantel, 2003), yet no consensus on either modifications to or a replacement for the traditional PERT/CPM planning and control technique have resulted.

As most research has been conducted in the single project environment, most critiques of PERT/CPM have also been focused on the single project environment. Kerzner (1994) states that PERT: (1) is end-item oriented—it separates the planners from the doers; (2) assumes infinite capacity; and (3) fails to recognize the lack of history on which to base estimates. Other researchers have found similar problems and have criticized certain PERT/CPM characteristics. Wiest and Levy (1977) question (1) whether activities themselves and their durations (and associated distributions) and precedent relationships can be known in advance; (2) the lack of cyclical and conditional activities; and (3) the assumption of an inverse linear relationship between cost and duration (activity crashing). Van Slyke (1963) and later Schonberger (1981) found that activity variability causes project duration to exceed PERT estimates, that is, as activity duration variability increases so does the difference between planned and actual project duration. Both found that PERT assumes path independence and questioned whether variance on one path might cause another path to be “late.” Van Slyke further identified the cause as interdependence of activities on the “independent” paths. Whether explicitly stated or simply suggested by the focus of their research, many researchers ultimately question the PERT/CPM assumption of infinite capacity and PERT/CPM’s disregard of activity duration variability.

Multiple Project Management Literature

Not unlike other business environments, the management of multiple projects has certain problems that must be recognized prior to the development of new tools for planning and control. Recent research in the area of multiple project planning and control has recognized several shortcomings of the PERT/CPM method. Researchers have explored resource assignment rules to better plan multiple projects (e.g., Lee et al., 1978; Trypia, 1980; Kurtulus and Davis, 1982; Kurtulus, 1985; Kurtulus and Narula, 1985; Allam, 1988; Mohanty and Siddiq, 1989; Bock and Patterson, 1990; Deckro et al., 1991; Dean et al., 1992; Abdel-Hamid, 1993; Kim and Leachman, 1993; Lawrence and Morton, 1993; Speranza and Vercellis, 1993; Yang and Sum, 1993; Vercellis, 1994; Tsai and Chiu, 1996) and have investigated the issue of multiple project control on both an organizational basis (e.g., Coulter, 1990; Platje and Seidel, 1993; Payne, 1995) and a tactical basis (e.g., Tsubakitani and Deckro, 1990; Dumond and Dumond, 1993). With the exception of Dumond and Dumond (1993) and Tsubakitani and Deckro (1990), this research has examined a static multiple project environment.

The investigations into the planning, scheduling, and control functions of multiple projects have found several fundamental characteristics inherent in multiple projects:

1. Multiple projects are interdependent due to the use of common resources.
2. Some method must be used to prioritize the use of resources among multiple projects.
3. There is some trade-off between the utilization of resources and the on-time completion of individual

projects.

4. Whether organizational or tactical, a control mechanism must exist to reduce the variance between planned and actual project completion dates.

Development of Guidelines

The widespread use of project management techniques and the general failure of projects to meet time, budget, or specification targets beg the examination of the fundamentals of project planning, scheduling, and control from a systems perspective. In the ensuing section, 12 guidelines for project planning, scheduling, and control based on this systems perspective are developed. These guidelines are posed as a starting point in the development of a comprehensive solution to project planning, scheduling, and control. Without this systems perspective of focusing on core problems of project failures, a proposed solution (such as better communications) may create more problems (such as having more project meetings and producing more reports) than it solves. The 12 guidelines are listed in the shaded box and are further explained in the text that follows.

Guideline I: Recognize the differences between due-date projects and money-making projects. The network structure may be the same, but a project to make money is started as soon as possible (to make money) and a project that is due by a given date is started as late as possible (to save money) while still providing protection for its completion. The project must be viewed as part of the larger system—what are the goal and objectives of the project with respect to the organization’s goal and objectives?

Guideline II: Recognize all of the activities required to achieve the goal of a project and the organization. In application, the goal of the project is generally a milestone in a much larger system. Ensure that the project scope fully defines the activities necessary to achieve the project goal and is in line with the system (organization) goal.

Guideline III: Recognize that 100-percent resource utilization may be counter to the objectives of the project and the organization goal. Plan resource use within and across projects such that the project is completed on time, on budget, and to full specifications.

Guideline IV: The rules for constructing project activity times must be known and practiced by all resources, resource managers, and project managers. A.5 probability of completion for the activities and project is required to determine a correct network. Padding (or buffering) should be applied strategically at the project level.

Guideline V: Minimize the amount of multitasking by critical resources and the amount of multitasking on activities on the critical path of the project to reduce activity lateness. Use multitasking cautiously—understanding its impact on project completion. Strategically buffer non-critical paths, resource contention, and project completions to reduce the impact of Murphy's Law (Murphy).

Guideline VI: Develop and implement a methodology for prioritizing resource allocation within a project and across projects so that resources know what is most important from the organizational (system) perspective.

Guideline VII: The project manager must consider all activities and dependencies to be completed to achieve the project goals, as well as all conditions that must be met before an activity can begin when developing the project network.

Guideline VIII: Recognize the existence of finite capacity and activity duration variability by changing the planning, scheduling, and control of single and multiple projects to

include a buffer time at the end of each individual project, as well as at points of convergence (technological convergence and convergence caused by resource contention) both within and across projects.

Guideline IX: Recognize that the current practice of minimizing costs by delaying activity expenditures might be counter to the objective of on-time delivery of the projects.

Guideline X: Recognize that measuring resource managers by resource utilization creates inflated activity time estimates, timing issues in the use of resources, multitasking within and across projects, and ultimately project lateness.

Guideline XI: Theory (research) must be revised to reflect and support practice. Activities on the Critical Chain or a near-Critical Chain should be scheduled to completion of the preceding activity instead of time. For example, traditional project management software usually provides schedules for resources and resources are available based on the schedule. Project research simulations initiate succeeding activities based on completion of preceding activities. At a minimum, research should define its parameters from a systems perspective to match reality.

Guideline XII: Establish a clear and effective method for the planning and control of multiple projects looking at resource contention across projects. Recognize that not all projects can start as soon as possible. Projects should be pipelined based on the capacity of critical resources and staggered based on the capacity of those resources.

Macro Issues

Goals, Objectives, and Measures of a Project

Granger describes the hierarchy of objectives as follows: “There are objectives with objectives, within objectives. They all require painstaking definition and close analysis if they are to be useful separately and profitable as a whole” (1964, 63). Remember, the goal of a project is to complete a project successfully, which usually translates into the lower level objectives of minimizing the costs associated with the project, completing the project on time, and completing the project as described in the specifications.

But what is the *real* goal of the project? Project goals can be classified as two types: (1) projects that have to be completed by a given date, and (2) projects that when completed generate revenues and therefore should be completed as soon as possible. The first type of project should be started as late as possible and still guarantee delivery of the project by the desired date (this strategy saves money). The second type of project should be started as soon as possible and guarantee delivery of the project by the desired date (this strategy makes money). The second type of project is much more commonplace, yet it is often treated as if it were the first type.

- Cause: PERT/CPM does not overtly recognize the difference between a project that must be completed by a promised due date (due-date project) and a project that must be completed in order to make money as quickly as possible (money-making project).

This is addressed by Guideline I.

Defining the Scope of a Project

In his research, Feldman (2001) identified the seven deadly sins of project estimating. Included are that not all tasks or

costs are defined and that estimates do not represent scope completeness. What is the goal of the project with respect to the goal of the organization (system)? What then is the scope of a project to support these organization and project goals? If we are defining a project that entails the opening of a new car dealership (see Rydell Group, 1995 for details of this example), what are the activities that should be included in the project? What is the goal of the project? In this instance, it is to make money from the sale of cars. In the past, a dealership would decide to open a new franchise in a large town. This might have been executed as if it were composed of several projects: buying the land; soliciting bids; breaking ground, constructing the facility, furnishing the building, and contracting utilities; ordering inventories; hiring, staffing, and training employees; etc. It normally takes nine months from the beginning of the project until the doors of the facility are open—nine months of money flowing out of the firm. By using the systems perspective and recognizing that the goal is to make money on the project (i.e., the completion of the project is to have a money-making machine in place producing), they might restructure the network as a single project by running several activities in parallel—building, ordering furnishings, setting up utilities, etc.—and complete the project in three months. The project end is defined as opening the door for business instead of completing construction. This recognition of the type of project and its real scope allows the business to generate profits after three months instead of after nine months.

- Cause: As traditionally applied, PERT/CPM does not recognize all of the activities required to achieve the goal of a project and the organization.

This is addressed by Guideline II.

The Project Management Dilemma

Williams (2001) states that scheduling presents a dual challenge in project management. On one hand, there is the need to match capacity to the demand placed on it—the resources must be available for use. On the other hand, idle capacity will cost the firm money—the resources must be fully utilized. The overall goal of most organizations is to be successful (to make more money in most cases), and all managers in the organization share this goal. A project manager is usually assigned to plan and execute a project successfully (on time, on budget, and to specifications). To ensure that the resources are available at the proper times, the project managers develop detailed schedules for each resource manager.

These resources, however, are not controlled by the project manager but are managed by resource managers. For the organization to be successful, the resource managers are charged with minimizing the operating expenses associated with using the resources. This directive to the resource managers usually translates into keeping the resources busy at all times. The resource managers are given a budget and measured against the budget and the efficient use of their resources. The dilemma is then between “The resource managers must have the resources available for the projects” and “The resource managers must keep their resources busy.” There is a constant tug of war between project managers trying to complete projects within time, budget, and specifications and resource managers trying to make efficient use of resources.

In a multi-project environment, the resource managers being pulled across activities on different projects exacerbates this dilemma; each project manager believing

that his or her project is the most important, most late, most critical, etc. In practice, the environment worsens. A product line manager might have a number of projects underway as does another product line manager. In a given time period, we can have a few projects in each product line competing with other projects in the product line or projects from different product lines competing for common resources. The resource manager is pulled from one activity to another, possibly without completing the first activity. The resource manager usually responds to the squeaky wheel, the project manager who yells the loudest, instead of having a formal mechanism for prioritizing activities within and across projects. In most situations, the resources also multitask (start one activity, stop, start another, stop, start another, stop, go back to the first, stop, go to the third, etc.) across activities, which generally extends the activity and completion times of each activity significantly and possibly delays the completion of one or more projects. This multitasking undoubtedly affects project quality as well.

- Cause: The resource and project managers are unable to effectively plan resource use across activities in the same project or across activities in different projects.
- Cause: Project managers are measured by their ability to meet their three objectives—complete projects on time, on budget, and to full specifications—whereas resource managers have the objective of utilizing their resources efficiently and are measured by their ability to keep resources fully utilized. These are conflicting objectives and have conflicting measures.

These are addressed by Guideline III.

Determining an Activity Time Estimate

In theory, we assume that the activity time is the mean time of the beta distribution (Miller, 1962). In reality, what time does the resource manager normally give? Is it the mean time? Seldom. Usually if a resource or resource manager is asked to provide a time estimate for a task, he or she pads the time a little (or a lot). If he or she ever misses that activity time and is chewed out by the project manager, then that time becomes more inflated to ensure successful activity completion. Think about it. If you provided a 50 percent task completion time estimate to your boss and were above the mean time 50 percent of the time, then your boss would probably think you were a poor worker. What probability of completion time do you give your boss? A completion time related to 50 percent or a time related to 95 percent? What if you finished the work early? Would you tell the project manager? You would probably not as his or her expectation would be that you should be able to complete future tasks in that amount of time. Yet, you gave the project manager a 95 percent probability of completion time to cover yourself. You would assume that if the project manager knew you finished early he would assume you provided inflated activity times and then would begin to question the time and costs you provided for other activities. Remember, we typically think the cost of a resource is based on the amount of time used by the resource to complete the job. If you consistently finished earlier than your time estimate, then the project manager would think you overpriced your resource. There is a strong tendency to both expand the time estimates of activities and, if the activity is finished early, not report the early finish.

Additionally, the project manager has a tendency to pad the project duration to ensure completion. Do you think the project manager is going to provide the boss a project completion time estimate that he or she is only 50 percent

sure of completing? He or she probably gives a 95 percent probability of completion as well. You only have to be late one time on a major project to learn to pad your project times. What does the overall manager then do with your project time estimate? Cut the project time and cost and expect the same specifications. Why? The project manager started as a resource, then worked as a resource manager, then worked as a project manager, and is now a general manager and has practiced and knows the rules of the game.

Frequently, resources must be used to work on more than one activity at a time. Why? Two conditions come to mind. The first condition is the practice of multitasking discussed previously. The second condition exists where the resource runs into an unanticipated delay (interestingly the delay could be caused by a missing activity or a missing technological arrow on the network; hence, the need to use a beta distribution with pessimistic times) or must set aside the activity until later. This condition is discussed in the next section on identifying obstacles to completing an activity.

- Cause: The rules and measures for determining activity time are ambiguous. For example, according to the assumptions of PERT theory, the resource (or resource manager) must provide a .5 probability time for activity times to build an accurate project network, yet the project manager expects a 1.00 probability of completion of the activity and project. If the resource finishes early or the project finishes early, then the expectation is for all of that resource's activities or project manager's projects to finish early.

This is addressed by Guideline IV.

- Cause: The resource and project managers are unable to plan resource use across activities in the same project or across activities in different projects.
- Cause: Murphy has struck and some activities have been delayed into the time allocated for other activities connected to the activity by technological precedence or by the use of a common resource (resource precedence).

These are addressed by Guidelines V and VI.

Development of a Project Network

In theory, the development of a project network is straightforward. First, we ask, “What are the activities of the project?” Then we ask, “What goes first? What goes next? What can be done in parallel?” These steps are over simplifications to say the least.

In reality, activities cannot start for a multitude of reasons not related to those activities preceding the delayed activity (e.g., the tools were not available, the materials didn’t come in from the vendor, the workforce wasn’t scheduled, the application was not made for a needed permit, etc.). In practice, each activity (node) in the network must be examined to determine what has to be present to perform the activity. The mere completion of the previous activities does not always constitute the starting condition for the activity. For example, in a recent upgrading of a computer network for a building, the work crew was scheduled, the users were notified that the building would be closed on the weekend, the computer was scheduled to be down, police were notified of workers crawling through ceiling spaces, etc., but the necessary cable did not arrive. Therefore, we find that while several activities were planned, one missing activity could delay the completion of the project and this

activity was not anticipated. There are many penetration points in a project that if something is not present, the activity (or activities) and possibly the project will be delayed. We have no means or warning of such situations in the traditional project management literature—we do use a beta distribution and place 1/6 probability of this pessimistic activity duration occurring. We need to re-examine the steps used in constructing a project network to reduce the likelihood of pessimistic activity times occurring. We need to fail-safe the activities.

For example, in Critical Chain project network development (using *Theory of Constraints* [TOC]), network developers use a prerequisite tree to identify obstacles to achieving each intermediate objective (activity). They ask, “What is preventing us from starting this activity?” Numerous obstacles are identified; in most cases, these are items not included on the original network. Then an activity to overcome this obstacle is identified and included in the network. In this manner, the network developers identify and include many “assumed activities” in the project and many connecting arrows (dependencies) that were omitted in the original network. Most networks created in this manner have at least 25 percent more activities (nodes) and are 50 percent denser (more arrows). A quick review of the causes of project failure from the literature review of 40 years of failures shows: techniques of estimating are poorly developed (project completion estimates are usually optimistic); too detailed or too broad activity structure; lack of planning; ineffective scheduling; critical tasks are left off the project plan; and, again, poor planning. All of these “causes of project failure” can be caused by activities (nodes) and missing dependencies (arrows).

A project network should include all of the activities and dependencies required to achieve the goal of the project—legal requirements, purchasing, designing, production, accounting, finance, marketing, sales, personnel, etc. Most

networks are used for the design and development stages and do not take a systems perspective to a project. The consequences are that the project may be completed in time (that time shown on the network), but the result of the project (making money or using the end product) is not achieved.

- Cause: The project network is not developed to include all obstacles that must be overcome before an activity might begin.

This is addressed by Guideline VII.

Micro Issues

The micro issues relate to errors or shortcomings in using the project tools. We will use simple numerical examples of each problem. By careful study of these errors and their causes, a systems approach to project planning, scheduling, and control can be developed and tested to ensure it addresses each of these problems.

Gedanken exercises, or thought experiments, have traditionally been used in the sciences rather than in business. The method uses logic and simple mathematics to construct an illustrative example to validate a hypothesis. While the method has usually been applied to scientific research areas such as quantum mechanics or astral physics where time, space, or both separate the subjects of scrutiny from the researchers, gedanken exercises also have the advantage of holding all other variables constant in order that the effects of the variable being examined are isolated. This simplification allows the researcher to gain knowledge and understanding by examining fragments of the system one piece at a time rather than losing the effects

of an individual variable in the noise of many interacting variables. With a full understanding of the behavior of each variable acting in isolation, the researcher might be able to construct a logically sound theory about the system.

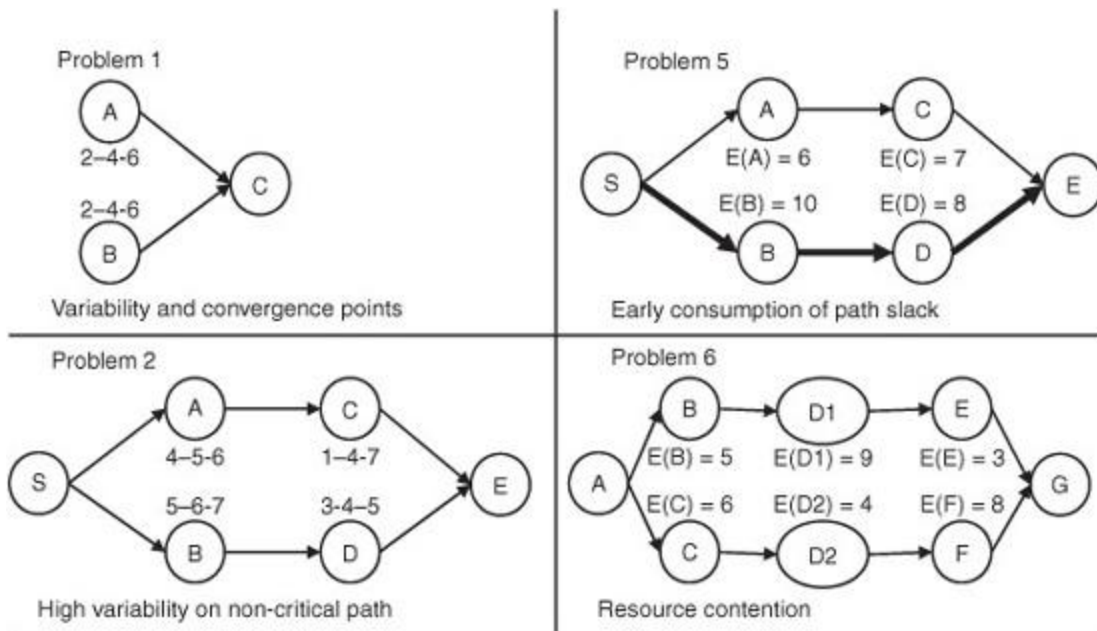
The use of gedankens in this research is based upon the realization that there are many factors that contribute to the project completion delays found in project planning, scheduling, and control. Here, the use of gedankens allows for the examination of each factor in isolation so that the factor can be studied to determine its effects on project completion.

Single Project Gedankens

Problem 1: Variability and Convergence Points The first of the eight weaknesses attributable to the assumptions of PERT/CPM is that of variability of activity duration and points of convergence. Many, if not all, PERT/CPM networks have points where two (or more) activities must be completed before a third activity may begin. Assume activity times follow a beta distribution. In [Fig. 2-3](#) Problem 1, activities A and B must be completed before activity C can begin. Since the expected duration of both A and B is 4 periods [$E(A \text{ or } B) = (2 + 4 \times 4 + 6)/6 = 4$], typical PERT/CPM planning would calculate that C will begin in period 4. However, if all possible combinations of the durations of both A and B are enumerated, the expected completion date of both A and B is 4.56 periods. The ultimate cause of the delay of activity C is the intersection of activities A and B (convergent point) when activity duration variability exists. With statistical fluctuations, convergent point calculations of start and finish times are incorrect.

- Cause: Network conventions require that all paths converge to one end node.
- Cause: Projects consist of dependent sequential activities, parallel paths, and convergent points.
- Cause: Murphy exists.
- Cause: PERT/CPM does not protect against Murphy.

These are addressed by Guideline VIII.



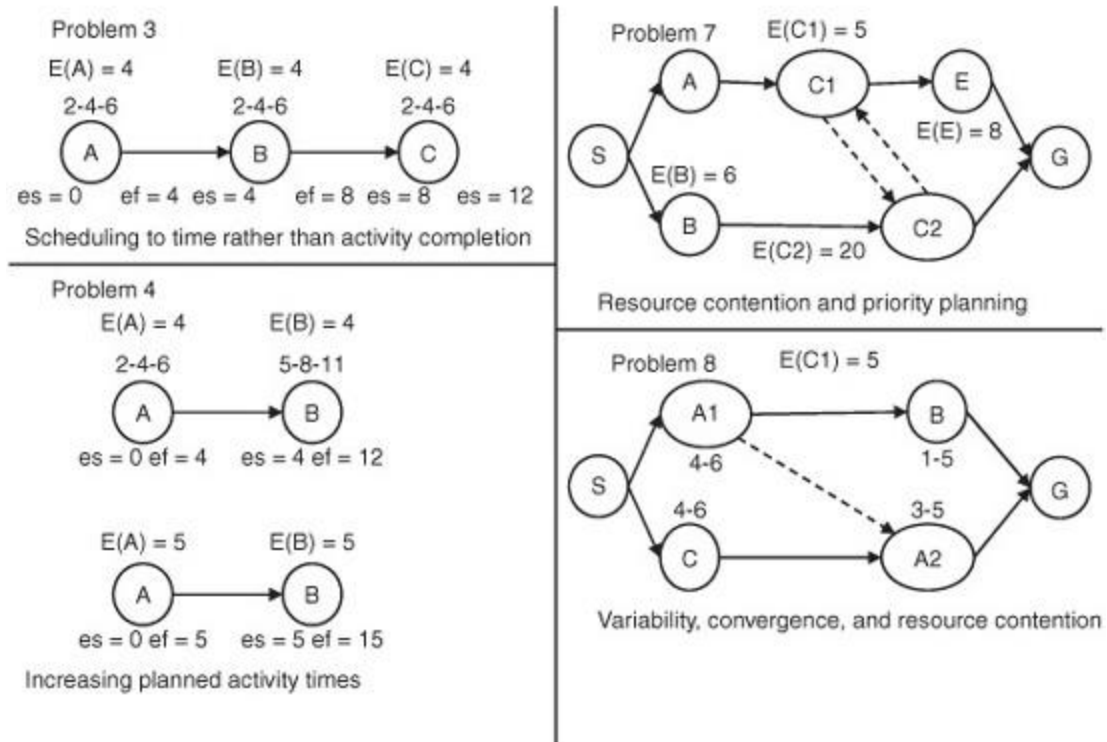


FIGURE 2-3 Eight problems with PERT/CPM management of single projects identified by Pittman (1994).

Problem 2: High Variability on a Non-Critical Path

[Figure 2-3](#) Problem 2 shows a simple PERT/CPM project with two paths. In typical PERT/CPM project management, the expected duration of each activity (assuming a beta distribution) is a simple point estimate based on the optimistic, most likely, and pessimistic estimates of activity duration. The expected duration of an activity is given by $E(A)$ of an activity.

$$E(A) = (5 + 4 \times 6 + 7) / 6 = 6$$

$$E(B) = (3 + 4 \times 4 + 5) / 6 = 4$$

$$E(C) = (4 + 4 \times 5 + 6) / 6 = 5$$

$$E(D) = (1 + 4 \times 4 + 7) / 6 = 4$$

The upper path (C-D) would be expected to take 9 periods, and the lower path (A-B) would be expected to take 10 periods. The PERT/CPM critical path is therefore A-B taking 10 periods. However, when all possible durations of each activity are enumerated, the expected duration of the project is not 10 periods but rather it is 10.725 periods. Van Slyke (1963) and later Schonberger (1981) suggested that near-critical paths be managed to ensure that variability on these paths does not affect the critical path. It is interesting to note that had the two paths not converged, the variability of the non-critical path would not have affected the critical path; therefore, this is a special type of convergence problem. (The assumptions of PERT include that a project has only one exit node; therefore, any project with more than one path must have a point of convergence.) The ultimate cause of the delay of project completion is the intersection of a non-critical path (C-D) with the critical path (A-B) when high activity duration variability exists on the non-critical path.

- Cause: Murphy exists.
- Cause: PERT/CPM does not protect against Murphy.

These are addressed by Guideline VIII.

Problem 3: Scheduling to Time Rather Than the Completion of the Prior Activity The managerial practice of scheduling to time rather than the completion of the prior activity is also affected by activity duration variability. [Figure 2-3](#) Problem 3 shows a simple three-activity PERT/CPM network. In practice, the typical PERT/CPM project manager generates and distributes to the resource managers a written or computer-generated schedule of planned activity

start times for that resource manager's resource based upon the expected duration of the preceding activities. Given that the expected durations of activities A, B, and C are 4, 4, and 4, respectively, a typical PERT/CPM schedule would be as follows:

Scheduled Activity	Start Date	Expected Duration	Finish Date
A	0	4	4
B	4	4	8
C	8	4	12

If, however, all possible combinations of activity duration are enumerated and the activities are started on the scheduled start date (or later if the preceding activity has not been completed), the project will have an actual expected duration of 13.11 periods. Project managers fail to take advantage of favorable completion times when the project is managed according to the above schedule. It should be noted here that optimistic completion times are leveraged only by the last activity in the network since there are no other activities planned to follow it. This means that the managerial practice of scheduling to time rather than completion of the previous activity is magnified in larger projects. The core driver, activity duration variability, is a fact. The practice of traditional project management to schedule to time instead of completion of preceding activity eliminates the opportunity to take advantage of optimistic completions of activities and thus produces poor project results.

- Cause: PERT/CPM does not recognize that some resources might be required for more than one activity.
- Cause: PERT/CPM provides resource schedules based only on technological relationships and time estimates.

These are addressed by Guideline XI.

Problem 4: Increasing Planned Activity Times

Resource managers (as opposed to project managers) have long felt pressure to complete their activity within the expected activity duration estimate. Resource managers then often increase the estimate of activity duration that they submit to the project manager to ensure that the activity is completed on time and to show high utilization of their resources. Low utilization of resources translates into excess resources that will be trimmed. [Figure 2-3](#) Problem 4 shows a simple two-activity PERT/CPM project. The upper project is the network that would be developed if the resource managers were to submit estimates of activity duration based on the actual estimates of activity duration. In the upper network, the PERT/CPM expected project duration would be 12 periods. The lower network is the PERT/CPM project network that the project manager would construct if each resource manager were to increase the expected duration of his activity by 25 percent.

Since the project manager constructs the project schedule based on the activity duration estimates provided by the resource managers, the resulting schedule would be as shown. The expected duration of the project would be 15 periods.

Scheduled Activity	Start Date	Expected Duration	Finish Date
A	0	5	5
B	5	10	15

If the project manager schedules to time (a resource schedule) rather than the completion of the prior activity, the actual expected duration of the project is 13.33 periods. In this case, the project manager receives praise for completing the project ahead of schedule and the activity managers receive praise for completing their respective

activities ahead of schedule (although only 67 percent of the time). If the estimates of activity duration had not been increased and the project manager had planned to time, the actual expected duration of the project would have been 12.67 periods. Clearly, this is a better result than 13.33 periods in both duration and cost, but the project manager would be punished for failing to meet the scheduled completion date. Finally, if the estimates of activity duration had not been increased and the project manager had scheduled to completion of the prior activity, the actual expected duration of the project would have been 12 periods. Again, the ultimate causes of project delay are resource managers including local protection in activity times and the project management practice of scheduling activity start times based on the expected time estimates instead of scheduling activities to start based on the actual completion of the preceding activity when variability exists.

- Cause: Murphy exists.
- Cause: Resource managers are expected to finish activities when planned.
- Cause: Resource managers do what they feel is necessary to ensure resource utilization and that the resources are available when promised.

These are addressed by Guidelines IV and X.

Problem 5: Early Consumption of Path Slack [Figure 2-3](#)

Problem 5 shows a simple PERT/CPM network. There are two paths in the network: A-C-E taking 28 periods and B-D-E taking 33 periods. The slack associated with the non-critical path is therefore 5 periods. Since activity E is critical (0 slack), all of the slack associated with the non-critical path can be “assigned” to activities A and C. Because a non-critical path has slack, a typical PERT/CPM project manager

would assign a start date of 5 to activity A. The expected finish date for activity C would therefore be period 18. If one examines the portion of the critical path before activity E (namely B-D), it is obvious that the expected finish date for path B-D is also period 18. It should be clear from the example of variability and convergence points that if activity duration variability exists in this network, then activity E cannot be expected to start in period 18. Consequently, the actual expected duration of the project cannot be 33 periods, but rather it must be longer. Two problems exist in the practice of project management. First, all of the slack associated with the non-critical path was absorbed in the planning stage of the project. PERT/CPM treats path slack as if it is associated with a specific activity and provides little recognition that once consumed by early activities, it is not available for protection for later activity (it is called activity slack, not path slack). Second, the project is delayed because of scheduling activity start times based on the PERT/CPM-calculated late start date rather than scheduling activities to start based on the actual completion of the preceding activity when variability exists.

- Cause: Murphy exists.
- Cause: The projects is a major undertaking, which determines the success or profitability (goal) of the organization.
- Cause: Project managers delay expenses by starting activities as late as possible.

These are addressed by Guidelines I, II, III, and IX.

Problem 6: Resource Contention Many researchers have recognized that the PERT/CPM assumption of infinite capacity does not accurately reflect the reality of finite capacity (e.g., Davis, 1966; 1973; Westney, 1991; Badiru,

1992; Davis et al., 1992; Dean, Denzler, and Watkins, 1992; Pittman, 1994; Zhan, 1994). When resource capacity is finite, the possibility exists that a single resource might be required to perform two or more activities simultaneously. Pittman defines resource contention as “the simultaneous demand for a common resource within a narrow time-span” (1994, 54).

[Figure 2-3](#) Problem 6 shows a simple PERT/CPM project with eight activities and two paths. In this example, variability of activity duration is ignored and only the expected activity duration estimate is used. The letter on the node designates the use of resources. There are only seven resources used to complete the eight activities. Resource D is used twice—once on node D1 and again on node D2. Typical PERT/CPM planning concludes that the lower path A-C-D2-F-G is the critical path, taking 30 periods, and the upper path A-B-D1-E-G is non-critical with 1 period of slack.

By examining the network, one can clearly see that resource D is required by activity D1 and activity D2 in period 8. Since resource D can only be used on one activity at a time, the activities must compete for the use of a limited resource. Either activity D1 uses resource D or activity D2 uses resource D, but both cannot use resource D simultaneously. By scheduling D1 and then D2 on resource D or vice versa, the duration of the project will be extended beyond 30 periods. The ultimate cause of project delay is the failure of PERT/CPM to recognize resource contention when resources are scarce.

- Cause: PERT/CPM does not recognize that some resources might be required for more than one activity.
- Cause: Resource utilizations are performance measures important to the organization’s success.

These are addressed by Guidelines III and VIII.

Problem 7: Resource Contention and Priority Planning It should now be clear that the PERT/CPM assumption of infinite capacity extends project duration when resource contention and limited resources exist. [Figure 2-3](#) Problem 7 demonstrates the effect on project duration of priority planning to overcome resource contention. The network shown has five activities and four resources. Once again, activity duration variability is ignored, and only the expected activity duration estimates are used. Typical PERT/CPM planning concludes that the lower path B-C2 is the critical path, taking 26 periods to complete, and the upper path A-C1-D is the non-critical path with 3 periods of associated slack.

If all activities are started on the early start date, the problem of resource contention occurs in period 15. If activity C2 is scheduled to use resource C first, then activity C1 must wait for the completion of activity C2 in period 26 before C1 can begin. In this case, the upper path and thus the project will not be completed until period 39. Conversely, if activity C1 is scheduled to use resource C first, then activity C2 must wait for the completion of activity C1 in period 15. In this case, the lower path and thus the project will not be completed until period 35. In either case, the duration of the project is greatly extended, but the difference between the two scheduling choices is not insignificant. The ultimate cause of project delay is the failure of PERT/CPM to provide a heuristic to prioritize resource use among activities when resource contention and limited resources exist.

- Cause: Priority of resource use may affect on-time project completion.

- Cause: PERT/CPM does not recognize that some resources might be required for more than one activity.
- Cause: PERT/CPM does not provide priority rules to support project completion.

These are addressed by Guidelines VI and VIII.

Problem 8: Variability, Convergence, and Resource Contention Activity duration variability can compound the problem of resource contention. In [Fig. 2-3](#) Problem 8, a simple PERT/CPM four activity, two-path project network is shown. There are only three resources required. If a uniform distribution of activity duration estimates is assumed, then the expected duration of each activity is as follows: $E(A1) = 5$, $E(B) = 3$, $E(C) = 5$, and $E(A2) = 4$. Typical PERT/CPM calculations conclude that resource contention does not exist since the expected completion date of activity A1 is period 5 and the early start date of activity A2 is also period 5. The lower path C-A2 is the PERT/CPM critical path, taking 9 periods. However, if activity A1 takes 6 periods to complete (a 50 percent probability), then a resource contention problem occurs, causing activity A2 to start later than its early start date and thus extending the project duration. If all possible combinations of activity duration are enumerated, the actual overall project duration is 9.75 periods. Activity duration variability causes resource contention when activity A1 requires 6 periods and causes a convergence problem when activity A1 and activity B require the longer of their respective estimates of duration. The cause of project delay is the failure of PERT/CPM to recognize convergence points, and resource contention and limited resources when activity duration variability exists.

- Cause: Network conventions require that all paths converge to one end node.

- Cause: Projects consist of dependent sequential activities, parallel paths, and convergent points.
- Cause: Murphy exists.
- Cause: PERT/CPM does not protect against Murphy.
- Cause: PERT/CPM does not recognize that some resources might be required for more than one activity.
- Cause: PERT/CPM does not view activity slack strategically.

These are addressed by Guidelines III and VIII.

Multiple Project Gedankens

Problem 1: Resource Contention across Projects Many researchers have recognized that the PERT/CPM assumption of infinite capacity does not accurately reflect the reality of finite capacity (e.g., Davis, 1966, 1973; Westney, 1991; Davis et al., 1992; Dean et al., 1992; Dumond, 1992; Badiru, 1993; Kerzner, 1994; Pittman, 1994; Zhan, 1994). When resource capacity is finite, the possibility exists that a single resource might be required to perform two or more activities simultaneously. Recall Pittman defines resource contention as “the simultaneous demand for a common resource within a narrow time-span” (1994, 54).

[Figure 2-4](#) Problem 1 shows two independent projects diagramed as a single “mega-project.” This method has been suggested by numerous researchers (Lee et al., 1978; Kurtulus and Davis, 1982; Kurtulus, 1985; Kurtulus and Narula, 1985; Mohanty and Siddiq, 1989; Bock and Patterson, 1990; Tsubakitani and Deckro, 1990; Deckro et al., 1991; Kim and Leachman, 1993; Lawrence and Morton, 1993; Yang and Sum, 1993; Vercellis, 1994), although there has been considerable debate over how to schedule resources.

The activity duration for each of the six activities in [Fig. 2-3](#) Problem 1 is deterministic (i.e., there is no variability). Activities B1 and B2 require the use of the same resource. Since there is only one of each type of resource and both activity B1 and activity B2 require the use of resource 2 in periods 7 through 15, a resource contention problem exists across the two projects. If resource contention is ignored as in typical PERT/CPM planning, project 1 has a planned completion date of period 17; project 2 has a planned completion date of period 18. There are two possible orderings of the use of resource 2—B1 then B2 and B2 then B1. If the project manager runs B2 then B1, project 2 would have the same completion date as typical PERT/CPM planning would estimate, but the completion date of project 1 would be delayed while activity B1 waits for activity B2 to finish using resource 2. If the project manager runs B1 then B2, activity B2 must wait for activity B1 to finish using resource 2, thus extending the completion of project 2. PERT/CPM does not provide mechanisms for determining how to optimally sequence activities on common resources across projects to provide realistic project completion times.

- Cause: PERT/CPM does not recognize that some resources might be required for more than one activity across projects.

This is addressed by Guidelines VIII and XII.

Problem 2: Priority of Resource Use across Projects

[Figure 2-4](#) Problem 2 shows two simple projects diagrammed as a single mega-project. Resource 3 is used by activities C1, C2, and C3. Typical PERT/CPM planning calculates the critical path for project 1 to be 23 periods and the critical path for project 2 to be 35 periods. There are three possible orderings for the use of resource 3 by the three activities:

C2-C1-C3 (solution 1 designated S1); C1-C3-C2 (solution 2 designated S2); and, C1-C2-C3 (solution 3 designated S3). Any of the three possible solutions—S1, S2, or S3—will delay the completion of the critical path of at least one of the two projects. In fact, solution 3 (C1-C2-C3) will delay the completion of both projects. Additionally, one can imagine the effects of multitasking or job splitting. Although PERT/CPM assumes that an activity, once started, cannot be stopped and restarted, it is common in practice for resource managers to do just that in order to appease various project managers. PERT/CPM does not provide guidelines on when and how to multitask, a common practice in industry.

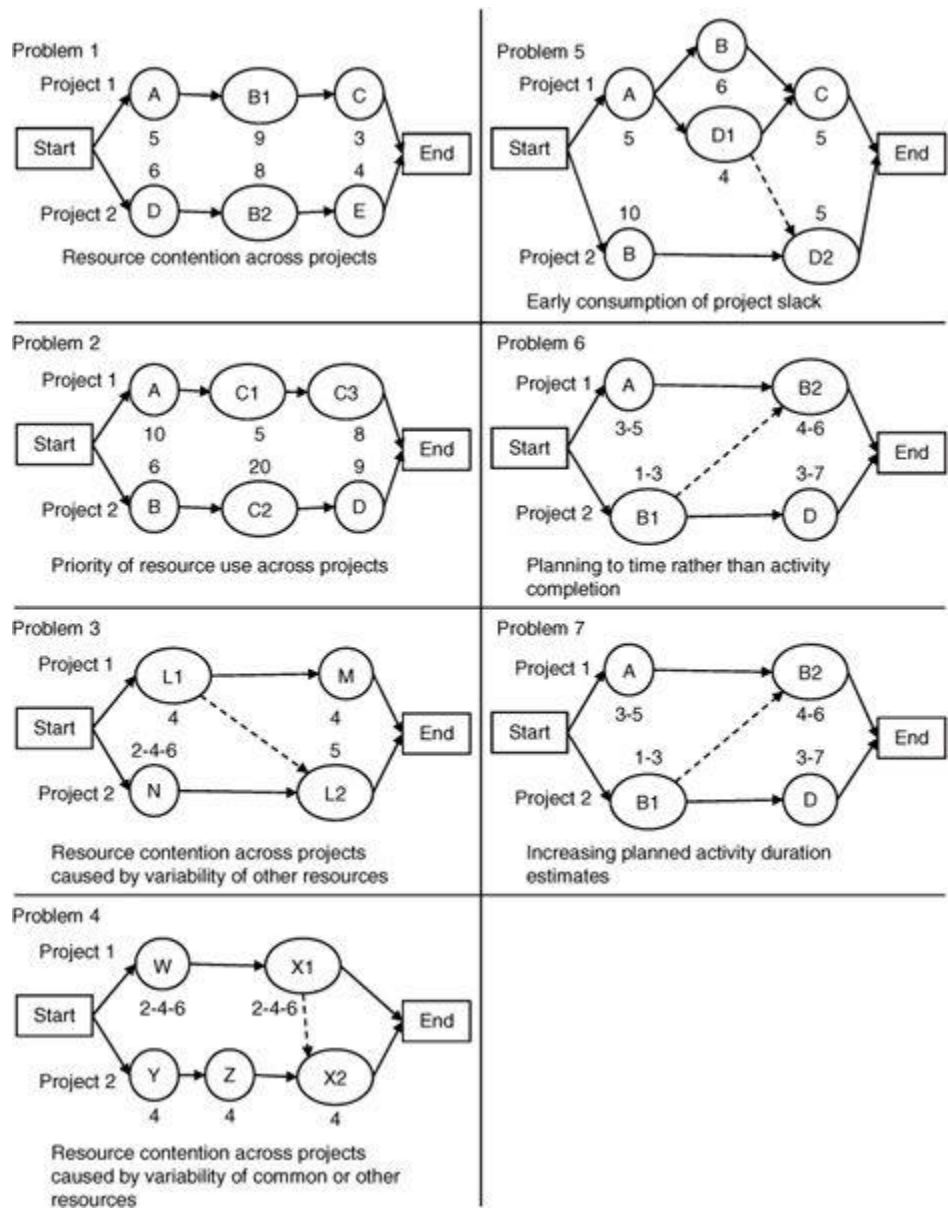


FIGURE 2-4 Seven problems with PERT/CPM management of single projects identified by Walker (1998).

- Cause: Priority of resource use across projects may affect on-time project completion.
- Cause: PERT/CPM does not provide priority rules to support project completion.

These are addressed by Guidelines VI, VIII, and XII.

Problem 3: Resource Contention across Projects Caused by Variability of Other Resources [Figure 2-4](#)

Problem 3 shows two simple projects diagrammed as a single mega-project. Each project has only two activities. Only one activity (activity N in project 2) has any associated variability. Resource 1 is used by activity L1 in project 1 and by activity L2 in project 2 in immediate succession. Typical PERT/CPM calculations estimate the completion date of project 1 to be period 8, and the completion date of project 2 to be period 8. The dashed arrow in [Fig. 2-4](#) Problem 3 shows the order of use of resource 1 as activity L1 then activity L2.

If all possible combinations of activity duration are enumerated, the completion date of project 1 is unaffected by the activity duration variability. However, when activity duration variability results in shorter than expected duration of activity N, a resource contention problem between activity L1 and activity L2 causes the delay of the completion date of project 2. Resource 1 is still in use by activity L1 when activity N is completed in period 2 (its optimistic estimate). As activity L2 cannot begin until activity L1 is completed, project 2 is unable to take advantage of an optimistic completion. A resource contention problem does not exist when activity N is completed in its pessimistic estimate. Therefore, only the late (pessimistic) duration time is added to the enumerated total, and the completion date of project 2 is later than planned. PERT/CPM does not recognize the impact of statistical fluctuation and dependent events on project completion. It should provide guidelines on buffering resources and paths against statistical fluctuations.

- Cause: Murphy exists.

- Cause: PERT/CPM does not protect against Murphy.
- Cause: PERT/CPM does not recognize that multiple projects are interrelated due to the shared use of common resources.

These are addressed by Guidelines VI, VIII, and XII.

Problem 4: Resource Contention across Projects Caused by Variability of Common or Other Resources

In [Fig. 2-4](#) Problem 4, two simple projects are diagrammed as a single mega-project. The activity durations in project 1 are variable, while the activity durations in project 2 are deterministic. Resource X is used by activity X1 in project 1 and by activity X2 in project 2 in immediate succession. Typical PERT/CPM calculations estimate the completion date of project 1 to be period 8, and the completion date of project 2 to be period 12. The dashed arrow in [Fig. 2-4](#) Problem 4 shows the order of use of resource 1 as activity X1 then activity X2.

If all possible combinations of activity durations for the two projects are enumerated, the completion of project 1 is unaffected by the activity duration variability of activities W and X1. However, when activity duration variability results in longer than expected duration of project 1, a resource contention problem between activity X1 and activity X2 causes the completion date of project 2 to be later than planned. PERT/CPM does not recognize the existence of these three core drivers and does not provide a mechanism to reduce their collective impact on project completion.

- Cause: Murphy exists.
- Cause: PERT/CPM does not protect against Murphy.
- Cause: PERT/CPM does not recognize that multiple projects are interrelated due to the shared use of common resources.

These are addressed by Guidelines VI, VIII, and XII.

Problem 5: Early Consumption of Project Slack [Figure 2-4](#) Problem 5 shows two simple projects diagrammed as a single mega-project. The critical paths of each project are as follows: project 1 A-B-C = 16; project 2 E-D2 = 15. The non-critical path of project 1 (A-D1-C) has two periods of associated slack. Typical PERT/CPM management would delay starting activity D1 in project 1 by the amount of slack available. If activity D1 is started on its late start date, activity D2 in project 2 and thus the completion of project 2 will be delayed by one period. PERT/CPM does not look at the impact of contention across projects on project lateness.

- Cause: PERT/CPM does not view activity slack strategically.
- Cause: The project is a major undertaking that determines the success or profitability (goal) of the organization.
- Cause: Project managers delay expenses by starting activities as late as possible.

These are addressed by Guidelines I, II, III, IX, and XII.

Problem 6: Planning to Time Rather Than Activity Completion [Figure 2-4](#) Problem 6 shows two simple projects diagrammed as a single mega-project. Each of the two projects has only two activities, and each activity has some associated activity duration variability. The expected duration of each activity is as follows: $E(A) = 4$, $E(B1) = 2$, $E(B2) = 5$, and $E(D) = 5$. Typical PERT/CPM planning would yield the following estimates of the date of project completion: project 1 complete in period 9, and project 2 complete in period 7.

The typical PERT/CPM manager would try to plan to start each activity based on the estimated time of completion of the preceding activity. Since $E(A) = 4$ and $E(B1) = 2$, the manager would plan to start activities B2 and D in periods 4 and 2, respectively. If all possible activity durations are enumerated and activities B2 and D are started based on the expected time of completion of A and B1, respectively, the expected completion dates of project 1 and project 2 exceed their PERT/CPM-planned completion dates. The expected completion date of project 1 is 9.5 periods versus 9; the expected completion date of project 2 is 7.625 periods versus 7. PERT/CPM does not look at the impact of scheduling to time instead of completion of preceding activity across projects on project completion.

- Cause: PERT/CPM does not recognize that some resources might be required for more than one activity.
- Cause: PERT/CPM provides resource schedules based only on technological relationships and time estimates.

These are addressed by Guidelines XI and XII.

Problem 7: Increasing Planned Activity Duration Estimates In this case ([Fig. 2-4](#) Problem 7), the activity duration estimates have been increased by one period to reflect that managers may recognize that activity duration variability exists (see [Fig. 2-3](#)). The revised estimates of activity duration are as follows: $E(A) = 5$, $E(B2) = 3$, $E(B1) = 6$, and $E(D) = 6$. If the reader examines [Fig. 2-3](#) Problem 4, he will find that increasing activity duration estimates leads to PERT/CPM-planned projects being late; increasing planned activity times in the multiple project environment also causes projects to be late.

If all possible combinations of activity duration are enumerated and activities B2 and D are started based on

the (revised) expected time of completion of A and B1, respectively, the expected completion date of project 1 is equal to its planned completion date given the revised estimates of activity duration. The expected completion date of project 2 is less than its planned completion date given the revised estimates of activity duration. The probability of on-time project completion is 100 percent for project 1 but only 50 percent for project 2. The augmenting (or “fudging”) of activity duration estimates has bettered the probability of on-time completion of project 1 (over [Fig. 2-3](#) Problem 6) and has not worsened the probability of on-time completion of project 2.

However, augmenting the activity duration estimates has caused the planned completion date of each project to be later than would be the case without increasing activity duration estimates. Both of the individual project completions have been extended for minimal or no gain in probability of on-time completion.

In this case, the project manager receives praise for completing the projects ahead of schedule and the activity managers receive praise for completing their respective activities ahead of schedule (though only 67 percent of the time). If the estimates of activity duration had not been increased and the project manager had planned to time (as in [Fig. 2-4](#) Problem 6), the expected completion date of project 1 would have been 9.5 periods and the expected completion date of project 2 would have been 7.625 periods. Clearly, this is a better result than 10 periods and 8 periods (projects 1 and 2, respectively) both in time and in cost, but the project manager would be punished for failing to meet the planned completion date. Additionally, had the activity duration estimates not been increased and the activities had been planned to completion, the expected completion dates of projects 1 and 2 would have been the same as their respective planned completion dates.

PERT/CPM does not discuss the impact of overestimating activity times across projects on project completion.

- Cause: Murphy exists.
- Cause: Resource managers are expected to finish activities when planned.
- Cause: Resource managers do what they feel is necessary to ensure resource utilization and that the resources are available when promised.

These are addressed by Guidelines IV, X, and XII.

The Use of PERT/CPM Critical Paths in the Single Project Environment

In conducting research on project management using simulation as compared to using project management in practice, some differences are noted. In most simulation models, the succeeding activity is linked to the completion of the preceding activity (scheduling by activity completion). For example, if activity A was scheduled to finish at time 10 but finished early at time 7, then activity B started at time 7 instead of waiting until the scheduled start time of 10. This seems to be the common practice in conducting research on project management.

In practice, however, with the growing use of project management software (Krakow, 1985; Lowery and Stover, 2001), the convention is to schedule by time (scheduling by time). Each resource is given a project schedule indicating when the resource is to start a given activity and how long it is to last. Where software is not used, either convention applies. Seldom, however, can a resource immediately reschedule what it is doing to start an activity early unless given warning.

The point here is that research does not simulate reality in its simplest form. In practice, where projects are large, several functions are involved, and project management software is used, projects seldom benefit from optimistic (early) completions. This means that the project is assured of being late unless extraordinary actions are taken to keep the project on schedule. If activities are only completed in their mean and pessimistic times, then activity times and project times are consistently understated. The project will always be late.¹

- Cause: Theory does not support practice.

This is addressed by Guideline XI.

The Use of PERT/CPM Critical Paths in the Multiple Project Environment

Two approaches are recommended in the research—the use of single project critical path and the use of a mega-project network connecting all projects to plan and controlling all projects simultaneously. Little research has been conducted to determine which of these is the better approach. Given the errors in logic of simulating projects as described previously, any research comparing these approaches needs to be reconsidered. Clearly, if resource contention exists across projects, then this must be reconciled to determine appropriate critical paths for each project and clearly, if one or a few resources are heavily loaded in most projects, then a mega-project approach is desired to ensure effective use of the constraining resource across projects.

In practice, 90 percent of projects occur in a multi-project environment and little research has been conducted in this environment. In practice, few organizations use the project networks to control projects and little research has been

conducted on how to control across multiple projects. After the original project plans are established, few bother with constantly updating the plans and rescheduling in the computer. Given all of the causes of failures of projects, one can see why a manager may not go to the trouble to constantly update every delay on every activity in a network.

- Cause: No well-defined method of planning and controlling projects in a multi-project environment exists.

This is addressed by Guideline XII.

Summary of the Micro Issues

That which is of the most importance here is not that researchers have not recognized that PERT/CPM is limited by its assumptions, but rather that the effects of these assumptions have been both underestimated and unstated. Many researchers have indeed recognized these assumptions, but there has been no systematic effort to eliminate the effects. By examining the gedankens, the reader will recognize that if the practitioner is forced to commit even one of the errors identified previously, then the project is probably going to be late. Additionally, the magnitude of the system effect (late, over budget, or under completed projects) is increased with each problem and each occurrence of each problem.

A Brief Overview of Critical Chain Project Management

Critical Chain in the Single Project Environment

Goldratt (1997) introduced the concept of *Critical Chain Project Management for Single Projects* (CCPM-SP or Critical Chain) to begin to address the problems associated with the more traditional methods of PERT/CPM and Gantt charts. As presented later, CCPM-SP addresses many of the guidelines listed previously, but not all of them. Guideline I concerns recognition of project type. Guidelines II and VII deal with the development of the project network, while Guideline XII is concerned with multiple projects. The resulting project network is a feasible, but not necessarily optimal, project plan.

[Figure 2-2](#) shows a typical activity-on-node PERT/CPM project network. Realistically, the completion of project activities requires the use of resources. Furthermore, resources are typically limited—there are only X number of programmers, or bulldozers, or whatever. Assume that the project shown in [Fig. 2-2](#) is to be completed using three different resources. [Figure 2-5](#) shows the same AON network as [Fig. 2-2](#), but with the addition of resources. The shading on the diagram denotes resource use: A and B, C and D, and E and F share a common resource.

The reader will quickly note that the activity times have been reduced by 50 percent. This reduction, at least partially, addresses Guidelines IV, VIII, and X. In addition, there has been another arrow added to the diagram. This dashed arrow represented the priority of resource use within the project—addressing Guidelines III and VI. By using the PERT/CPM technique of forward- and backward-passes through the network considering this newly added dashed arrow, the ES/EF and LS/LF times can be determined. Those activities with zero slack are critical activities. However, the sequence of activities (A-D-C-F) does not correspond to a PERT/CPM path, so the term *chain* is used to denote the

difference between a PERT/CPM path (which considers only technological precedence) and a CCPM-SP chain (which considers both technological and resource precedence). Since all of the activities on the chain A-D-C-F have zero slack, this chain is called the *Critical Chain (CC)*.

Other additions to the diagram are the boxes labeled FB and PCB. These boxes denote feeding buffers and the project completion buffer, respectively. These buffers exist to address Guidelines V, VIII, and IX. Time was taken out of each of the activities in the project, resulting in a .5 probability that each activity will be completed on time. The buffers exist to increase the probability of on-time project completion. The PCB adds time to the end of the project. In this case, since the CC is 20.5 days the PCB would be 10.25 days—the project can then be promised to be delivered in 30.75 days. The feeding buffers exist to protect the CC from variation of non-critical activities. If activities B and E were started on the LS date and suffered any delay, then the CC would be jeopardized. The feeding buffers require that these activities be started sometime before their LS date. (Actual determination of buffer size is left to later chapters.)

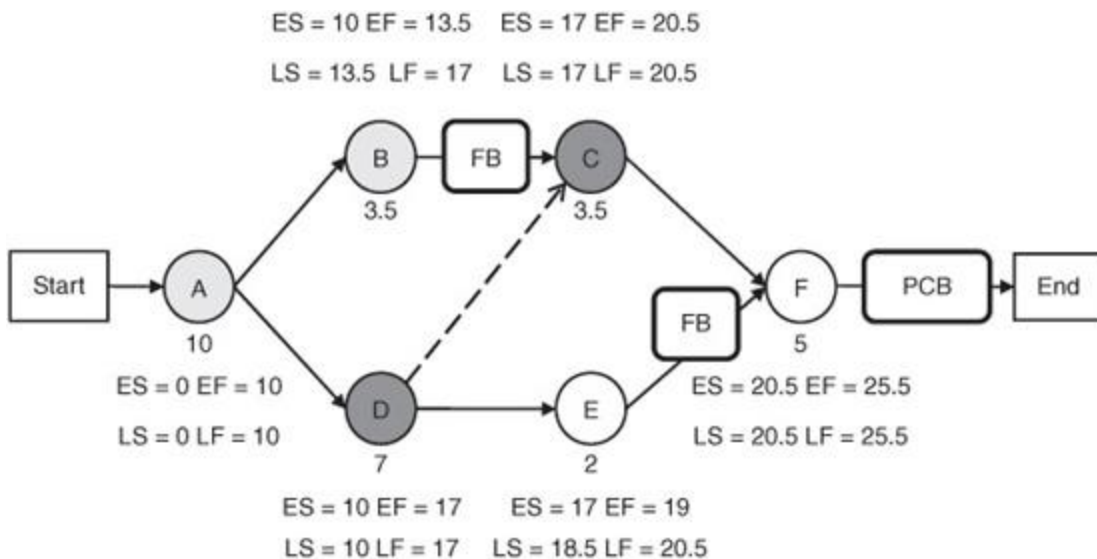


FIGURE 2-5 Typical activity-on-node project network with resource contention identified (shading shows same resource use).

In practice, CCPM-SP requires that all activities on the CC be monitored and started as soon as the previous activity ends in order to take advantage of early completions. This process addresses Guideline XI. Additionally, all resources on the CC are monitored to ensure that multitasking is eliminated or minimized to address Guideline V.

Brief Review of Critical Chain Literature

In the book *Critical Chain*, Goldratt (1997) first published the concept of CCPM. Like several of his prior texts, the book outlined the concept in a narrative fashion and does not seem to have been intended to be a “how-to” manual for CCPM. Rather, its purpose seems to have been to provide a basis for a stream of research that might be pursued by him and others. Pittman (1994) and Walker (1998) examined the single and multiple project environments (respectively) and sought to expose the assumptions and practice of scheduling and controlling projects by traditional methods. Their work provides the basis for the *gedankens* presented earlier in this chapter.

Hoel and Taylor (1999) sought to provide a method (via simulation) for determining the appropriate size for the buffers required by CCPM. Rand (2000) introduced CCPM to the project management literature framing CCPM as an extension of TOC. He concluded that CCPM not only dealt with the technical aspects of project management (like PERT/CPM) but also that CCPM dealt with how senior management manages human behavior in the construction of the project network as well as the execution of the network. Steyn (2000) followed this research with an

investigation of the fundamentals of CCPM. He concluded that a major impediment to implementing CCPM is that it requires a fundamental change in the way project management is approached and that such a change is likely to meet with resistance.

However, Herroelen and Leus (2001) argued that while CCPM was as important to project management as TOC was to production scheduling, CCPM oversimplified the issue of scheduling and rescheduling. Herroelen, Leus, and Demeulemeester (2002) continued much of the same argument in a later paper. Likewise, Raz, Divr, and Barnes re-examined CCPM and concluded that project performance is often a function of the skills and capabilities of project leaders and that “some CCPM principles do make sense in certain situations” (2003, 31). McKay and Morton (1998) as well as Pinto (1999) were concerned that CCPM might be misapplied by managers who failed to understand the underpinnings of CCPM and who attempted to adopt it without fully changing their fundamental approach to the management of projects.

Answering this criticism, Steyn (2002) sought to apply TOC to a variety of other areas of project management beyond the creation and execution of project schedules. He recognized the multidisciplinary nature of project management and how it affects cash flow, stakeholder needs, and risk management. Yeo and Ning (2002) began work on integrating supply chain management with project management. Sonawane (2004) incorporated systems dynamics with CCPM to create a “modern” project management system. Similarly, Lee and Miller (2004) applied systems thinking to multiple projects along with CCPM, and Trietsch (2005) argued that CCPM is, in fact, a more holistic approach to project management than traditional methods.

Herroelen and Leus concede that CCPM “seems practical and well thought-out...nevertheless, for single projects, the

unconditional focus on a 'Critical Chain' seems useless..." (2004, 1616). Srinivasan, Best, and Chandrasekaren (2007) presented a case study that clearly appears to contradict this conclusion. The *Warner Robins Air Logistics Center* (WR-ALC) is charged with the repair and overhaul of C-5 transport aircraft. After an eight-month implementation period starting in 2005 and without the addition of any resources, WR-ALC returned five additional aircraft to the operational fleet by reducing the number of in-service planes from 12 to 7. The replacement value of these aircraft is \$2.4 billion and does not consider nonmonetary benefits such as increased responsiveness and casualty avoidance during wartime.

Summary and Conclusions

The literature of project management relates to the practice of project management and to the theory of project management. The practice emphasizes the large number of project failures, and the theory focuses on fine-tuning algorithms in an attempt to minimize computer time or project duration. Certainly, the two themes should converge to provide a simple but effective approach to project management. One approach to refocusing the theoretical literature is to take a different perspective to project management. Our systems approach attempts to specifically identify several of the sources of project failure.

The purposes of this chapter were twofold. First, we examined the macro issues associated with project management. Second, the micro issues of project networks were examined. The overall objective of this research is not to propose solutions to each of the surface problems revealed in this chapter but to identify and logically link these surface problems to their underlying causes as well as to project failure. One must fully understand the core

problems of project management and its environment before proposing a comprehensive solution to these core problems. Without this systems perspective, a proposed solution may create more problems than it solves.

This chapter provides evidence that, not unlike other business environments, the management of single and multiple projects has certain core problems that must be recognized prior to the development of new tools for planning and control. Recent research in the areas of both single and multiple project planning and control has recognized the shortcomings of the PERT/CPM method. Twelve guidelines have been proposed by which the effective planning and control of both single and multiple projects might be improved.

These guidelines reflect fundamental changes to the way single and multiple projects are currently planned and controlled. The objective of any improved planning and control technique should not be to find the optimal solution to each of the problems found in project management, but rather that a feasible or realistic solution be found for all of the problems in project management (such as that presented by Goldratt, 1997 and Newbold, 1999). Furthermore, solutions for any one problem should not be developed in isolation of the other problems. Additionally, complex solutions should not be developed for such solutions are difficult for practitioners to both understand and apply. A realistic planned project completion date that is met is better than an optimal solution that is never met.

The practitioners have voiced the strongest criticisms of current project management methods. This is because the current reward systems are based not upon the method used but upon the results received—on time, on budget, and to full specifications. Recognizing the inadequacies of PERT-based methods in achieving the desired results, practitioners attempt to modify these project management methods.

Practitioners recognize the effects of variability and finite capacity when their projects are completed late and over-budget, but do not understand the underlying reasons for the observed effects. They intuitively “know” that PERT/CPM assumptions are causing their project to fail, but have not recognized that their own behavior is also a cause. These behaviors, such as project managers seeking to delay spending or to avoid late penalties, or resource managers increasing planned activity duration to protect their resources, are driven by policies and measures. (A full discussion of policies and measures, how they influence behavior, and how to align individual behaviors with corporate goals is far beyond the scope of this chapter.)

Important recommendations to practitioners may be made because of the research presented in this chapter. Project managers should understand that estimates of activity duration are prone to over-estimation and, counter-intuitively, often lead to poor project performance. Project managers should also understand that multiple projects are interdependent due to the shared use of common resources. As such, decisions made with respect to one project may have detrimental effects on other projects, even projects that have not yet started. Additionally, a reward system should be developed that recognizes the completion rather than the duration of both activities and projects. Resource managers need to understand the concept of a Critical Chain and must also take advantage of early activity completions. Finally, project managers should not base planned project completion dates on PERT-based plans, but rather on some method that recognizes the shared use of common resources and the existence of statistical fluctuation within and across projects.

Researchers have identified many of these surface problems in their studies; however, no comprehensive examination of the causes of these surface problems has been undertaken. We feel this approach is not enough. To

provide a practical framework for reducing project failures, a systems approach must be taken to identify both macro and micro surface problems, core drivers (environmental factors), and core problems with the PERT/CPM methodology. We do not propose a comprehensive solution to addressing project management; however, we do provide some guidelines to start a dialog with other researchers in developing a more effective and practitioner-friendly approach to project management. Researchers should use these guidelines as a starting point to develop algorithms that are more robust. Goldratt's Critical Chain method offers promise in addressing many of these problems. It has been used effectively in a limited but growing number of different environments. That method and others need to be developed and refined to provide a systems perspective encompassing the needs of project managers, resource managers, and organization managers.

Policies, procedures, measures, planning, and control methods need to be re-examined as indicated by the current reality trees of single and multiple project organizations. Underlying conflicts among the goals and measures of managers create many of the surface problems seen in a project management environment. These conflicts must be resolved by providing supporting policies, procedures, and measures. Given that these can be devised and successfully implemented, a systems perspective must be utilized to identify all of the core drivers in a given environment and the planning and control system so structured to accommodate these core drivers. The project environment has several common core drivers that must be incorporated into any planning and control methodology. We tried to identify a number of these and provide guidelines for managers to consider in planning and controlling projects. These guidelines should also provide the foundation for further research into developing and testing

effective methodologies for planning and controlling projects.

Academia needs to shift emphasis from defining a good algorithm from one that minimizes computer time or finds the shortest completion time to determining ways to construct networks to guarantee completion of the project on the plan and methods of immunizing projects against statistical fluctuation. The recognition that in the presence of statistical fluctuation and dependent events, lateness accumulates are essential. Methods of eliminating or minimizing the effect of the accumulated lateness on project completion are needed. Strategic buffering of resources, paths, and networks in single and multiple projects must also be studied.

References

- Abdel-Hamid, T. K. 1993. "A multiproject perspective of single-project dynamics," *Journal of Systems and Software* 22(3):151-165.
- Allam, S. I. G. 1988. "Multi-project scheduling: a new categorization for heuristic scheduling rules in construction scheduling problems," *Construction Management and Economics* 6(2):93-115.
- Avots, I. 1970. "Why does project management fail?" *Management Review* 59(10):36-41.
- Badiru, A. B. 1992. "Critical resource diagram: A new tool for resource management," *Industrial Engineering* 24(10):58-59, 65.
- Badiru, A. B. 1993. "Activity-resource assignments using critical resource diagramming." *Project Management*

Journal 24(3):15-21.

Bildson, R. A. and Gillespie, J. R. 1962. "Critical Path Planning—PERT Integration", *Operations Research* 10(6):909-912.

Black, K. 1996. "Causes of project failure: A survey of professional engineers," *PM Network* 21-24.

Bock, D. B. and Patterson, J. H. 1990. "A comparison of due date setting, resource assignment, and job preemption heuristics for the multiproject scheduling problem," *Decision Sciences* 21(2):387-402.

Brooks, F. P. 1995. *The Mythical Man-Month*. Anniversary Edition. Boston: Addison-Wesley.

Brown, D. 2001. "Lack of skills to blame for project failures," *Canadian HR Reporter* 14(17): 1-12.

Clark, C. E. 1961. Comments on the Proceeding Paper (*The PERT Model for the Distribution of an Activity Time*). *Operations Research* 10(3):348.

Coulter III, C. 1990. "Multiproject management and control," *Cost Engineering* 32(10):19-24.

Davis, E. W. 1966. "Resource allocation in Project Network Models—A survey," *Journal of Industrial Engineering* 17(4):177-188.

Davis, E. W. 1973. "Project scheduling under resource constraints: Historical review and categorization of procedures," *AIIE Transactions* 5(4):297-313.

Davis, K. R., Stam, A., and Grzybowski, R. A. 1992. "Resource constrained project scheduling with

multiple objectives: A decision support approach," *Computers & Operational Research* 19(7):657-669.

Dean, B. V., Denzler, D. R., and Watkins, J. J. 1992. "Multiproject staff scheduling with variable resource constraints," *IEEE Transactions on Engineering Management* 39(1):59-72.

Deckro, R. F., Winkofsky, E. P., Hebert, J. E., and Gagon, R. 1991. "A decomposition approach to multi-project scheduling," *European Journal of Operational Research* 51(1):110-118.

Dumond, J. 1992. "In a multiresource environment: How much is enough?" *International Journal of Production Research* 30(2):395-410.

Dumond, E. J. and Dumond, J. 1993. "An examination of resourcing policies for the multi-resource problem," *International Journal of Operations Management* 13(5):54-76.

Feldman, J. I. 2001. "The seven deadly sins of project estimating," *Information Strategy* 18(1):30-36.

Goldratt, E. M. 1997. *Critical Chain*. Great Barrington, MA: North River Press.

Granger, C. H. 1964. "The hierarchy of objectives," *Harvard Business Review* 42(3):63-74.

Gutierrez, G. J. and Kouvelis, P. 1991. "Parkinson's law and its implications for project management," *Management Science* 17(8):990-1001.

Healy, T. L. 1961. "Activity subdivision and PERT probability statements," *Operations Research* 341-

348.

Herroelen, W. and Leus, R. 2001. "On the merits and pitfalls of critical chain scheduling," *Journal of Operations Management* 19(5):559-577.

Herroelen, W. and Leus, R. 2004. "Robust and reactive project scheduling: A review and classification of procedures," *International Journal of Production Research* 42(8):1599-1620.

Herroelen, W., Leus, R., and Demeulemeester, E. 2002. "Critical chain project scheduling: Do not oversimplify," *Project Management Journal* 33(4):49-60.

Hoel, K. and Taylor, S. G. 1999. "Quantifying buffers for project schedules," *Production and Inventory Management Journal* (40)2:43-47.

Hughes, M. W. 1986. "Why projects fail: The efforts of ignoring the obvious," *Industrial Engineering* 14-18.

James, G. 2000. "Beware of consultants peddling snake oil," *Computerworld* 34(39):40.

Kelley, J. E. 1962. "Critical-path planning and scheduling mathematical basis," *Operations Research* 296-320.

Kerzner, H. 1994. *Project Management: A Systems Approach to Planning, Scheduling, and Controlling*. 5th ed. New York: Van Nostrand Reinhold.

Kim, S. and Leachman, R. C. 1993. "Multi-project scheduling with explicit lateness costs," *IIE Transactions* 25(2):34-44.

- Klingel Jr., A. R. 1966. "Bias in PERT project completion time calculations for a real network," *Management Science* 13(4):194-201.
- Krakov, I. H. 1985. *Project Management with the IBM PC Using Microsoft Project, Harvard Project Manager, Visischedule, Project Scheduler*. Bowie, MD: Prady Communications Co.
- Krakovski, M. 1974. "PERT and Parkinson's law," *Interfaces* 5(1):35-40.
- Kurtulus, I. 1985. "Multiproject scheduling: Analysis of scheduling strategies under unequal delay penalties," *Journal of Operations Management* 5(3):291-307.
- Kurtulus, I. and Davis, E. W. 1982. "Multi-project scheduling: Categorization of heuristic rules performance," *Management Science* 28(2):161-172.
- Kurtulus, I. and Narula, S. C. 1985. "Multi-project scheduling: Analysis of project performance," *IIE Transactions* 17(1):58-66.
- Lawrence, S. R. and Morton, T. E. 1993. "Resource-constrained multi-project scheduling with tardy cost: Comparing myopic, bottleneck, and resource pricing heuristics," *European Journal of Operational Research* 64(2):168-187.
- Lee, B. and Miller, J. 2004. "Multi-project software engineering analysis using systems thinking," *Software Process Improvement and Practice* 9(3):173-214.
- Lee, S. M., Park, O. E., and Economides, S. C. 1978. "Resource planning for multiple projects," *Decision*

Sciences 9(1):49-67.

Levy, F. K., Thompson, G. L., and Wiest, J. D. 1962. "The ABCs of Critical Path Method," *Harvard Business Review* 98-108.

Lowery, G. and Stover, T. 2001. *Managing Projects with Microsoft Project 2000 for Windows*. New York: John Wiley & Sons.

Malcolm, D. G., Roseboom, J. H., and Clark, C. E. 1959. "Application of a technique for research and development program evaluation," *Operations Research* 646-669.

Marks, N. E. and Taylor, H. L. 1966. "CPM/PERT: A diagrammatic scheduling procedure," *Studies in Personnel and Management* (18). Austin: Bureau of Business Research, Graduate School of Business, University of Texas.

Matta, N. F. and Ashkenas, R. N. 2003. "Why good projects fail anyway," *Harvard Business Review* 109-114.

McKay, K. N. and Morton, T. E. 1998. "Critical chain," *IIE Transactions* 30(8):759-762.

Meredith, J. R. and Mantel, S. J. 2003. *Project Management: A Managerial Approach*. 5th ed. New York: John Wiley & Sons.

Middleton, C. J. 1967. "How to set up a project organization," *Harvard Business Review* 73-82.

Miller, R. W. 1962. "How to plan and control with PERT," *Harvard Business Review* 93-104.

- Millstein, H. S. 1961. "Comments on the proceeding paper (Healy)," *Operations Research* 349-350.
- Mohanty, R. P. and Siddiq, M. K. 1989. "Multiple projects—Multiple resources constrained scheduling: A multiobjective approach," *Engineering Costs & Production Economics* 18(1): 83-92.
- Neimat, T. 2005. "Why IT projects fail," The Project Perfect White Paper Collection.
<http://www.projectperfect.com.au>.
- Newbold, R. 1999. *Project Management in the Fast Lane: Applying the Theory of Constraints*. Boca Raton, FL: St. Lucie Press.
- Paige, H. W. 1963. "How PERT-cost helps the general manager," *Harvard Business Review* 87-95.
- Parkinson C. N. 1957. *Parkinson's Law and Other Studies in Administration*. New York: Random House.
- Payne, J. H. 1995. "Management of multiple simultaneous projects: A state-of-the-art review," *International Journal of Project Management* 13(3):163-168.
- Pinto, J. K. 1999. "Some constraints on the theory of constraints—Taking a critical look at the Critical Chain," *PM Network* 13(8):49-51.
- Pinto, J. K. and Mantel, S. J. 1990. "The causes of project failure," *IEEE Transactions on Engineering Management* 37(4):269-275.
- Pinto, J. K. and Presscott, D. P. 1988. "Project success: Definitions and measurement techniques," *Project*

Management Journal 19(1):67-71.

Pinto, J. K. and Slevin, D. P. 1987. "Critical factors in successful project implementation," *IEEE Transactions in Engineering Management* EM-34(1):22-27.

Pinto, J. K. and Slevin, D. P. 1989. "The project champion: Key to implementation success," *Project Management Journal* XX:15-20.

Pitagorsky, G. 2001. "A scientific approach to project management," *Machine Design* 73(14): 78-82.

Pittman, P. H. 1994. Project management: A more effective methodology for the planning and control of projects. Unpublished doctoral diss., University of Georgia.

Platje, A. and Seidel, H. 1993. "Breakthrough in multiproject management: How to escape the vicious circle of planning and control," *International Journal of Project Management* 11(4):209-213.

Pocock, J. W. 1962. "PERT as an analytical aid for program planning—Its payoff and problems," *Operations Research* 10(6):893-903.

Rand, G. K. 2000. "Critical chain: The theory of constraints applied to project management," *International Journal of Project Management* 18(3):173-177.

Raz, T., Divr, D., and Barnes, R. 2003. "A critical look at critical chain project management," *Project Management Journal* 34(4):24-32.

- Roseboom, J. H. 1961. "Comments on a paper by Thomas Healy," *Operations Research* 909-910.
- Rydell Group. 1995. TOC at Saturn and GM Dealers. Paper presented at the North American Jonah Upgrade Conference, September 21-24, Philadelphia, PA.
- Schonberger, R. J. 1981. "Why projects are 'always' late: A rationale based on manual simulation of a PERT/CPM network," *Interfaces* 11(5):66-70.
- Sonawane, R. 2004. Applying systems dynamics and critical chain methods to develop a modern construction project management system. Unpublished master thesis, Texas A&M University-Kingsville.
- Speranza, M. G. and Vercellis, C. 1993. "Hierarchical models for multi-project planning and scheduling," *European Journal of Operations Research* 64(2):312-325.
- Srinivasan, M. M., Best, W. D., and Chandrasekaran, S. 2007. "Warner Robins Air Logistics Center streamlines aircraft repair and overhaul." *Interfaces* 37(1):7-21.
- Steyn, H. 2000. "An investigation into the fundamentals of critical chain project scheduling," *International Journal of Project Management* 19(6):363-369.
- Steyn, H. 2002. "Project management applications of the theory of constraints beyond critical chain scheduling," *International Journal of Project Management* 20(1):75-80.
- Trietsch, D. 2005. "Why a critical path by any other name would smell less sweet? Towards a holistic

approach to PERT/CPM," *Project Management Journal* 36(1):27-36.

Trypia, M. N. 1980. "Cost minimization of simultaneous projects that require the same scarce resource," *European Journal of Operations Research* 5(4):235-238.

Tsai, D. M. and Chiu, H. N. 1996. "Two heuristics for scheduling multiple projects with resource constraints," *Construction Management and Economics* 14(4):325-340.

Tsubakitani, S. and Deckro, R. F. 1990. "A heuristic for multi-project scheduling with limited resources in the housing industry," *European Journal of Operational Research* 49(1):80-91.

Van Slyke, R. M. 1963, "Monte Carlo methods and the PERT problem", *Operations Research* 11(5):839-860.

Vercellis, C. 1994. "Constrained multi-project planning problems: A Lagrangean decomposition approach," *European Journal of Operational Research* 78(2):267-275.

Yang, K. and Sum, C. 1993. "A comparison of resource allocation and activity scheduling rules in a dynamic multi-project environment," *Journal of Operations Management* 11(2): 207-218.

Yeo, K. T. and Ning, J. H. 2002. "Integrating supply chain and critical chain concepts in engineer-procure-construct (EPC) projects," *International Journal of Project Management* 20(4):253-262.

- Walker II, E. D. 1998. Planning and controlling multiple, simultaneous, independent projects in a resource constrained environment. Unpublished doctoral diss., University of Georgia.
- Westney, R. E. 1991. "Resource scheduling—Is AI the answer?" 1991 *American Association of Cost Engineers Transactions* K.6.1-K.6.9.
- Wiest, J. D. and Levy, F. K. 1977. *A Management Guide to PERT/CPM with GERT/PDM/DCPM and Other Networks*. 2nd ed. Englewood Cliffs, NJ: Prentice Hall.
- Williams, D. 2001. "Right on time." *CA Magazine* 134(7):30-31.
- Zhan, J. 1994. "Heuristics for scheduling resource-constrained projects in MPM networks," *European Journal of Operational Research* 76(1):192-205.

About the Author

Ed D. Walker II, Associate Professor of Management at Valdosta State University, is from Milledgeville, Georgia. He is recognized as a CPIM by APICS and as a Jonah by the Avraham Y. Goldratt Institute and is certified in TOC project management, the TOC thinking processes, and TOC operations management by the Theory of Constraints International Certification Organization. He has a BS in Business Administration and Math/Physics from Presbyterian College and an MBA in Finance from Auburn University. Prior to receiving his PhD in Operations Management at the University of Georgia, Dr. Walker worked in production planning and control, distribution, and plant management in both the food processing and textile industries. He has published over 20 journal and conference articles in the

areas of Theory of Constraints, project management, manufacturing planning and control systems, performance measurement, and classroom pedagogy. Two young children keep Dr. Walker and his wife quite busy. He enjoys volunteering at his church, working outdoors, officiating high school football, as well as hunting and fishing.

CHAPTER 3

A Critical Chain Project Management Primer

Charlene Spoede Budd and Janice Cerveney

Introduction

As evidenced by their support of professional certification from the Project Management Institute,¹ organizations want to improve their project management skills. Even though the profession has recognized the need to improve and companies seriously try to improve their project management maturity, most are still on the lower levels of a typical five-level project management maturity model, and few have reached the top levels involving continuous improvement.

The previous chapter, by Ed Walker, is an excellent review of the entire history of project management. The next three chapters, one by Realization, one by Rob Newbold, and one by AGI, cover the latest *Critical Chain (CC)* advances. Compared to the other chapters in this section, this chapter contains tutorial material on how CC works, along with some implementation suggestions. Our basic assumption is that the reader knows little or nothing about *Critical Chain Project Management (CCPM)*.

Why These Widespread Project-Related Problems Persist

[Chapter 2](#) clearly outlines a host of very familiar problems with which *project managers (PMs)* continue to struggle. History suggests that a definitive solution is elusive.

Throughout his professional life, Eli Goldratt has stressed how complex and chaotic situations can be handled with a simple five-step approach (first detailed in Goldratt and Cox, 1984; Goldratt, 1990, 59-62). This same approach applies to project management (Leach, 2005, 52-54). The first of the five steps involves identifying the constraint. For projects, the constraint that prevents an organization from earning more, both now and in the future, is the time required to complete a project with available resources. In product development projects, for example, projects delivered late may lose a significant share of their potential market to competitors.

Copyright © 2010 by Charlene Spoede Budd and Janice Cerveney.

For traditionally-managed projects, two assumptions guarantee project completion delays: (1) project task times can be accurately predicted, and (2) the traditional project management planning and control system is effective (Leach, 2005, 10-11). Resources are asked to provide an estimate of the time required to complete a particular task. Once all project resources have reported their estimated (and safe) times, management frequently requires lower estimates. If those estimates are accepted by all resources (and resources usually have little choice), the estimate becomes a commitment upon which the resource will be evaluated.

Task Duration Uncertainty

We know that task times follow a distribution pattern that is skewed to the right. No task can be completed in zero time,

but the maximum possible time can be extremely long. Look at a simple example such as the time required to drive to an important client's office. Let's say if you pressed the speed limit (exceeded it by 5 or 6 miles—9 or 10 is more common in Atlanta) and encountered no problems, you might make the trip in 20 minutes (the minimum task time). Normally, however, the trip takes about 30 minutes. If there was an accident on the freeway that you couldn't avoid, it may take several hours. If you had to promise your client that you would be there at a certain time or lose your account, how much time would you estimate? It would certainly not be 20 or 30 minutes. The same is true for a project resource who must promise to complete a task in a certain amount of time. The estimate typically will be in a range such that the resource has a 90 to 95 percent probability of successful on-time completion.

Since task times follow a skewed distribution, as illustrated in [Fig. 3-1](#), and have unique properties, completion times cannot be estimated with precision. Nevertheless, an estimated time must be provided. Resources operating in traditional project management environments, therefore, are forced to protect their careers by providing times with appropriate safety that will permit them to survive management "adjustments" and to deliver on their promises.

The area under the curve shows the probability of completing the task in a given time estimate. Estimated completion time if resources could dedicate their time to the task, without interruption, most likely would occur somewhere to the left of the longer vertical dotted line in [Fig. 3-1](#) (between the two arrows pointing in opposite directions). A minimum time, the far left point in the distribution curve, can occur, but with very low probability. To provide for interruptions and urgent but unplanned assignments, resources typically elect to provide a time that they are 90 to 95 percent confident they can achieve. In

general, if resources deliver on the accepted due date, they receive a good evaluation. If a task is delivered late, their evaluation is diminished, depending on how late a task is delivered. Typically, resources are evaluated based on how well they perform their assignments, independent of other resources working on the same projects.

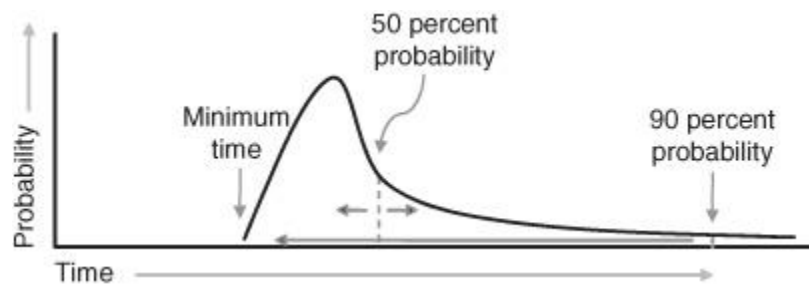


FIGURE 3-1 Probabilities for a task with a skewed distribution.

In project-based environments, where multiple projects are performed using shared resources, “accurate” task estimates are even more critical for planning achievable schedules. Because making sure that an individual resource is not assigned to work on two tasks at the same time is logistically next to impossible in a multi-project environment (due to task completion uncertainty, where no single point estimate can be correct), only the most sophisticated (project mature) organizations attempt to solve this *mathematically NP-hard problem*.²

The mechanisms used to plan and schedule projects must minimize the risk of nonproductive, abortive, or misdirected effort. The methodology also must provide relevant, timely information for management control, such that appropriate intervention occurs when needed during project execution. In addition, the system must capture the correct information for improvement.

In traditional multi-project environments, a basic problem is the inability to ensure adequate progress on the projects

already underway while simultaneously having the flexibility to take advantage of new business opportunities as they arise. Typically, new projects are entered into a system as soon as they are funded and few organizations appear to be able to successfully establish stable global project priorities.

Traditional Survivor Behaviors

Human resources may be assigned to three to five major projects, sometimes in addition to their functional duties. To deal with skewed task times, *official* resource estimates, those that are turned in to management, generally are two or more times their estimated *dedicated* durations. *Dedicated duration estimates are those that could be met if resources were allowed to work without interruption.* However, most project employees do not work without interruption.

Forced multitasking induces additional stress on already heavily loaded resources. In spite of the widespread praise for multitasking ability, most people realize that they are more productive when they concentrate their effort on one task (Rubenstein et al., 2001; Shellenbarger, 2003, 169).

Three behaviors typically are used by resources to deal with chaotic project situations: (1) student syndrome, (2) sandbagging, and (3) engaging in Parkinson's Law. They are discussed in the following sections.

Student Syndrome

The name *student syndrome* (Goldratt, 1997) developed from a common student behavior of lobbying for an extension to an exam date that is two weeks away (typically some time after an upcoming school event) so they can study. However, most students only begin studying for the exam a few hours or, at best, a couple of days before it is

scheduled—whether or not they receive the requested delay. While this behavior is typical for students, it also is typical for the rest of us.

Negotiating additional time would appear to enable us to ensure on-time completion of current assignments. Of course, when we wait until the last possible minute to begin a new assignment, we should expect that we would run into problems we had not anticipated. Therefore, meeting the promised due date may be extremely difficult and stressful.

Sandbagging Completed Work

Sandbagging refers to *holding completed work until a more beneficial time arrives to officially acknowledge its completion*. A resource may have fought long and hard for the time allotted to its task. Therefore, if a task is completed early, there may be a very real reluctance to pass it on to the next activity, since their next task duration estimate may be discounted accordingly. Also, acknowledging an early completion frequently results in additional assigned work, increasing a resource's workload even more. In order to protect one's reputation and believing that the next resource will not be prepared to take advantage of an early start if one discloses early completion, most experienced resources will not pass on their work until just prior to, or on, the due date.

Sales people (including those who sell projects) who have met their quotas regularly engage in sandbagging. A similar delay in passing on work, but due to a different motivation, is work completion delays due to Parkinson's Law, discussed next.

"Improving" a Completed Task

Rather than merely holding completed work, if work on a task proceeds extremely well (which normally would enable the work to be completed before the due date), there is a tendency among some resources to continue improving the completed work. This sometimes is referred to as “polishing” work and has come to be known as *Parkinson’s Law* which states that *work expands to fill the time available* (Parkinson, 1957).

Not infrequently, these resources think they are improving the quality of their product by adding “extras” not included in the original specifications for the task. (In our experience, this is especially true on software projects.) However, the unspecified and undocumented addition may cause problems, sometimes major problems, further along in the project.

Management rarely distinguishes between task uncertainty and the time that is lost when tasks are started late, constantly interrupted, or when workers fail to turn over finished work. CC acknowledges these dysfunctional behaviors and establishes policies to deter their occurrence. The next section summarizes the basic elements of CC.

Key Elements of Critical Chain

While many of the basic project management concepts are preserved in CCPM, it is designed to overcome the most egregious issues that have resulted in the poor performance of projects as described in the previous chapter and in all-too-familiar press reports. The magnitude of change required demands a different approach. When people are doing their best and outcomes are unacceptable, as Deming (1993, 172-175) so strongly advised, we must change the system.

Changes are required in planning, scheduling in single and multi-project environments, and in managing the project.

Issues in Creating a Project Plan

Most stakeholders involved in a project are quite familiar with the general requirements of the project that include issues such as identifying the project objective, having a project charter, understanding the work breakdown structure, acquiring resources, and creating a plan for the budget and scheduled tasks.³ Once planned, most project management books suggest that the *critical path*, the *longest chain of dependent tasks*, is the most important in project completion. Therefore, this path is given preferential treatment when assigning scarce resources.

When planning a CC project, the total budget may be the same, but there are particular scheduling requirements that differ from the traditional critical path approach. However, we will discuss the scheduling differences then return to the project budget toward the end of this chapter.

Task Duration Estimates

Human resources naturally include safety time in their duration estimates. In defining a CC schedule, this safety is removed from individual (local) tasks and aggregated to protect the entire project. It can be helpful if the PM has some historical knowledge of an individual resource's safety preferences. In general, about half of a task's "safe" time, the time required to be 90 to 95 percent confident of task completion, is there to cover interruptions, surprise rework, urgent unanticipated assignments, and task estimation error.

Rather than providing "start" and "finish" times for every task, as recommended by traditional project management, CC uses task durations and asks resources to work on a *first-in, first-out (FIFO)* basis for all queued tasks. Start times

are provided only for initial activities on a path—those with successor activities but no predecessors.

Task Uncertainty

Just as a management reserve is established to cover the uncertainty of estimated costs, task uncertainty is managed in CC with buffers of time. Besides referring to these blocks of time with no scheduled activities as buffers, some U.S. government guides call them *schedule reserves* or *schedule margins*.⁴ (For example, see NASA, 2009, 223–224; United States Government Accountability Office [GAO], 2009, 56, respectively, for NASA and GAO best-practices remarks).

Buffers will be explained more fully later and will be illustrated in an example of a project scheduled using CC concepts.

Resource Contention

In most traditional project plans, encountering resource unavailability or tasks delivered late can cause the critical path to shift. Some projects will have the critical path shift several times during project execution. These shifts result in constantly changing priorities and continuously revised task start and finish times. This is especially true if projects are not leveled prior to initiation of project work.

In CC project plans, it is vital to resolve all resource contention with reverse passes through the project schedule; that is, starting from the end of the project schedule, eliminating resource contention all the way back to the start of the project. Following this resource leveling effort, the Critical Chain is identified as the longest chain of task and resource dependencies. Ideally, a Critical Chain remains the same throughout project execution.

Merging Paths

There is special risk in a project schedule where paths or chains⁵ of dependent activities merge with other chains. If one of the paths is the Critical Chain, the project completion date can be endangered by late completion of a non-critical path. As we will see in a sample CC project schedule, special attention is ascribed to chains of dependent activities that merge into tasks on the Critical Chain.

Communications

There are many policy differences between traditional project management and CCPM and those differences will require changes in organizational and individual behaviors. An especially important process for CC projects is an effective communication system that includes a method of *resource notifications, a message to a resource to: (1) start a chain (path) of activities, (2) prepare for upcoming work on the Critical Chain, or (3) perform critical work on a higher-priority project in a multi-project environment.* Such notifications help to ensure that CC tasks, which determine project completion, will be given appropriate priority.

Later in the chapter, we will describe how CC overcomes all the forces that pose challenges to successful project completion.

Issues in Managing Project Execution

Ideally, no project should be started unless all specifications have been received, the charter has been approved, an acceptable schedule has been approved, and all other preparatory steps have been accomplished. Further, no task should be started unless all required materials are available and the task is at the start of a FIFO work queue. Having everything ready and on hand before starting a project or a

task is referred to as having a “whole kit” or a “full kit.” While a research project may violate this “rule,” other projects should not.

In traditional project management, once a project is begun, each task is managed as if it were an independent event. That is, a worker is rewarded if an assigned task is completed on or before its scheduled finish date; exhorted to work harder if it is not completed on the finish date; and punished, in several ways, if the finish date is overrun by a significant amount. The rationale for this partitioning of project work is that if every task is completed on time, the project will be completed on time. Of course, this rationale completely disregards the reality that few, if any, tasks are passed on early. Therefore, if only a few tasks complete late, as nearly always happens, the entire project is delayed.

Critical Chain uses buffers to manage task duration uncertainty and to monitor project progress. A later section, entitled “Project Control: The Power of Buffer Management,” describes how this is accomplished.

Scheduling a Single Project

One of the easiest ways to illustrate the way CC addresses the issues presented previously is by contrasting what is done in traditional project environments with a single project example. To illustrate the scheduling steps, we use a simple project with 5 resources and 10 activities (tasks). There is only one of each of the five uniquely qualified resources; each resource can perform its own work but cannot perform the work of any other resource.

To provide a better understanding of single-project scheduling, a manual CC process is described in the following section. However, there are software programs now available that can perform these scheduling steps in both single project and multi-project environments. The

advantages of the CC solution in a multi-project environment are even more dramatic and will be discussed later.

Modifying Task Duration Estimates

Following initial project planning activities (i.e., identifying the project objective, authorizing the project charter, determining required tasks and the work breakdown structure, etc.), a most critical step in preparing a project plan is getting estimates for task durations. In most organizations with a functional or *matrix structure*,⁶ project resources are primarily responsible to a line manager and only secondarily to a PM. The resources know that project tasks will be in addition to their usual job responsibilities. They often do not know how much of their time the task will take. They do know that they will be expected to complete their project tasks within the estimated (promised) time.

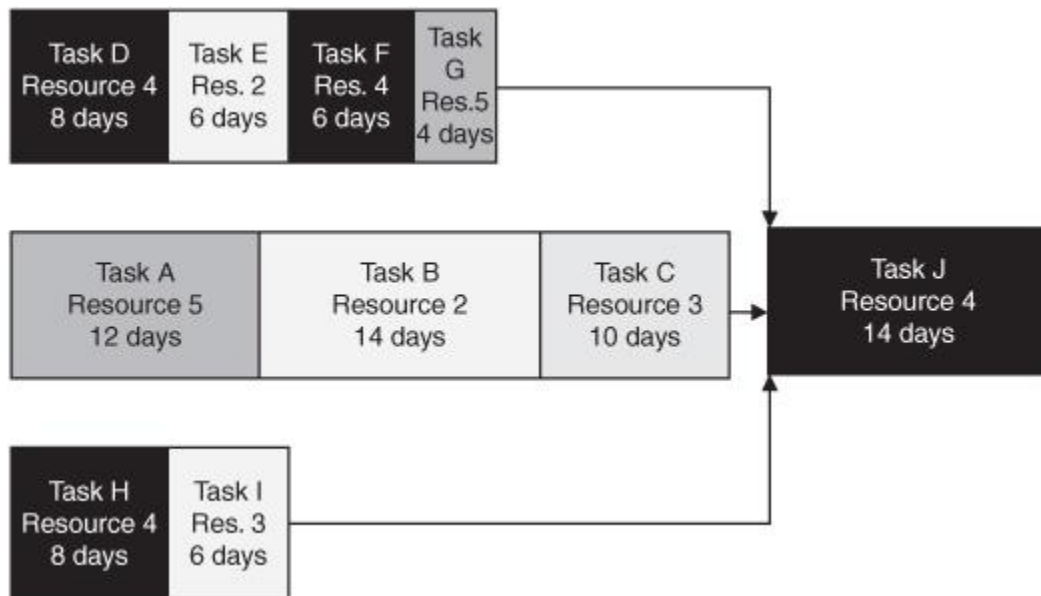


FIGURE 3-2 Estimated *dedicated* task times.

If resources could work uninterrupted on a task until it is completed, they probably would provide the estimated durations shown in [Fig. 3-2](#). However, they would be risking their jobs to report these durations to their PM.

Veteran resources (all of whom have experienced unplanned work assignments and interruptions that affect their ability to complete their assigned tasks on time) rely on their intuitive knowledge that the actual task time will be an element of a skewed distribution having some minimum time but possibly a very high maximum. Therefore, knowledgeable resources typically give a task duration estimate that they can expect to meet at least 90 percent of the time. (Recall that resources in traditional project management are held responsible for completing the task by their estimated times.)

Assume, for our simple example, resources provide the task times illustrated in [Fig. 3-3](#) for a traditional, resource-leveled project.⁷ Task D, on the top path of [Fig. 3-3](#) and Task J at the very end of the project, show the skewed continuous distributions associated with the task estimates of 16 days and 28 days for Tasks D and J, respectively. All of the tasks have similar distributions that justify the times submitted, although they are not shown in the figure. The tasks that comprise the critical path are highlighted with a solid thick grey line from Task D and two-thirds of Task E, to Task B, shifts back to Task E (to complete the final 4 days of the 12-day task) estimate, and then continues to Tasks F, G, and J.

Note that the lower path of activities in [Fig. 3-3](#) is scheduled to begin as soon as possible (immediately after Resource 4 completes Task D), as is the general practice in traditional project scheduling. The generally held erroneous assumption that an early start helps ensure an early finish guarantees path starts as soon as possible. The project is scheduled to complete in 104 workdays. Microsoft Project

2007™ software splits the work on Task E into two parts and includes Task B on the critical path, as reflected in [Fig. 3-3](#).

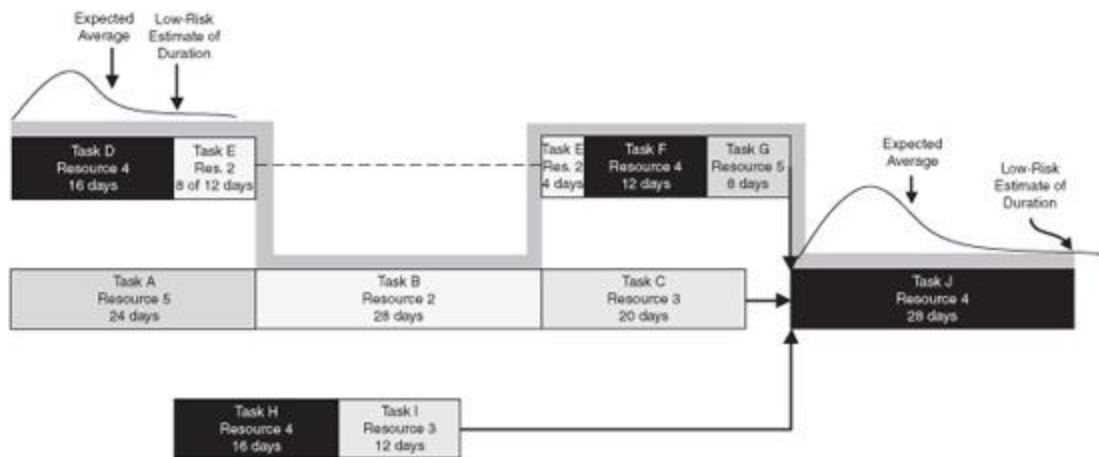


FIGURE 3-3 Traditional resource-leveled project schedule (showing 2 of 10 distributions).

A major precept of the *Theory of Constraints* (TOC) states that *the sum of local optima is not equal to global optima*. In managing a project, the concept implies that concentrating on individual task completion does not ensure that the project will complete on time. The entire project may be in danger of not completing on schedule when even a few tasks are late (especially if they are on the Critical Chain). This means that we should change our focus from individual task completion to project completion. This focus is accomplished in a CC schedule by removing the safety (time) built into the individual tasks and concentrating this safety where it will protect the project's completion rather than the completion of individual tasks.

Can we really do this and not jeopardize the completion of every task? Yes, but it will take some changes in organizational behavior patterns, which we will discuss later. First, let's look at the statistics that really indicate that there is little overall danger in removing some time from task duration estimates.

A Bit of Statistics

Basic statistical understanding informs us that about half a project's tasks will complete before their *dedicated* duration and about half will complete after. The uncertainty in the sum of tasks is equal to the square root of the sum of the squares of the individual task variations. Variation here is the difference between the estimated and actual time.

$$\sqrt{(\text{Difference in Task A})^2 + (\text{Difference in Task B})^2 + \dots + (\text{Difference in Task J})^2}$$

Of course, the above formula technically is only applicable in repetitive situations where task durations are independent, but it helps us understand a complex issue.

Intuitively, when we amass all the protection in one place (a buffer), the early and late finishes should offset each other. Thus, TOC argues that we need only about half the safety used to protect each individual task. For shorter projects, where the offsets might not happen as expected, we might need more than 50 percent of the safety removed from individual tasks; on larger projects, we may not need as much. However, 50 percent is a good rule of thumb for establishing a project's *buffer*, the schedule reserve that we establish at the end of the Critical Chain.

Critical Chain Scheduling

Armed with knowledge of CC issues and the single project environment, we are prepared to schedule the sample project introduced previously. There are six generic steps in CC scheduling

1. Build an initial project schedule that has safety times (assumed here as approximately 50 percent of the

original task time estimate) removed from task durations.⁸

2. Working from the end of the project, eliminate all resource contention (first backward pass).
3. Identify the longest path of resource and task dependencies: the Critical Chain (the second backward pass).
4. Calculate and insert the project buffer (typically about half the safety removed from tasks on the Critical Chain).
5. Calculate and insert feeding buffers for all paths (chains) merging into the Critical Chain, resolving any newly discovered resource contention within the project. (Compute buffer sizes using the same procedure as that for the project buffer.)
6. Add communication *resource buffers*⁹ to ensure timely notifications to resources that have no predecessors to begin work, and to all resources that have work assigned on the Critical Chain.

An optional seventh step may be required if the planned completion date is too far in the future.

7. Analyze the schedule and evaluate options to complete the project at an earlier date; make selected changes, review and approve changes, and update the schedule.

As we will see, for most CC projects it is easy to know which additional resources should be acquired, and for what periods of time.¹⁰ Therefore, we will concentrate on the first six steps in our sample illustration.

Critical Chain Scheduling—Steps 1 through 4

To schedule the project shown in [Fig. 3-3](#) as a CC project, the safety embedded in each task is removed from protecting the task (a local optimum) and half of the safety related to the CC tasks is moved to a place where it can protect the entire project from uncertainty. That means that the starting point for developing a CC schedule is shown in [Fig. 3-2](#), i.e. the project with the estimated dedicated task times.

Step 2, resource leveling, then is accomplished by starting at the end of the project and working backward, rescheduling or shifting each task so that there is no overlap of tasks assigned to the same resource while keeping the total length of the project as short as possible. Unlike the traditional approach where resource leveling occurs after identification of the critical path, with CC the leveling is accomplished prior to identification of the Critical Chain.

Step 3 involves another backward pass through the project in order to identify the most obvious candidate for the longest path. Once again, starting from the end of the project, and working backward on the chosen path, the Critical Chain (☆) is identified as Task J, C, and B, but then, because Task E uses the same resource as Task B, the Critical Chain moves up to Task E and finally Task D.¹¹

Step 4 results in the insertion of the project buffer. The size of the buffer, technically, is half the number of time units (days in this example) of safety that were removed from the activities that comprise the Critical Chain. Most practitioners understand that it is half the total length of the Critical Chain. The project buffer is placed at the end of the Critical Chain, thus pushing the end date past the apparent end point of the last task.

[Figure 3-4](#) demonstrates the first four steps in CC scheduling: (1) using shortened or dedicated task times, (2) resource leveling, (3) the identification of the Critical Chain, and (4) insertion of the project buffer. The Critical Chain is

identified with white stars beside the task identifications in [Fig. 3-4](#).

Note that the project buffer in [Fig. 3-4](#) (Step 4) has no task or resource assigned. The project buffer can be used to manage the time lost in those tasks that do not complete in their shortened (dedicated) time. Rather than having safety in individual tasks where it may not be required (and typically is wasted due to the student syndrome, sandbagging, and Parkinson’s Law), the project buffer protects completion of the project. Note also that we have rescheduled the lower chain of tasks in [Fig. 3-4](#) to start as late as possible without encountering resource contention on Tasks F and H.

The project is now scheduled to complete in 78 days, but there are several more steps. Remember that software is available to perform these steps.

In terms of the TOC Five Focusing Steps (5FS), the CCPM scheduling Steps 1, 2, 3, and 4 would be the TOC focusing Step 1 (identify the constraint) and Step 2 (exploit the constraint).

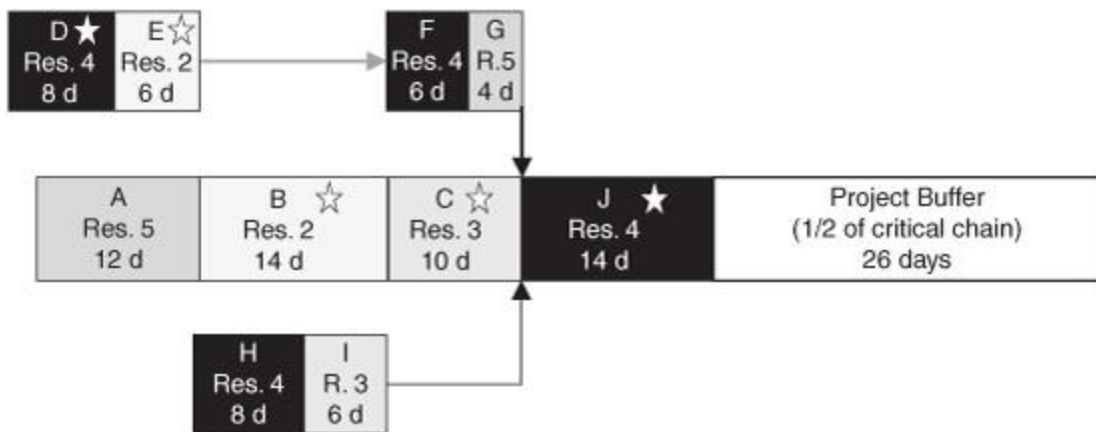


FIGURE 3-4 Critical Chain incomplete project schedule with only a project buffer.

Merging Paths—Step 5

When non-Critical Chains of dependent activities that merge into the Critical Chain encounter problems, the entire project can be delayed. To provide protection for such possibilities, feeding buffers, Step 5 should be added at the end of each non-critical path at the point where it joins the Critical Chain. Like the project buffer, feeding buffers are blocks of time that do not have assigned tasks or resources.

The size of these buffers is determined using the same logic as with the project buffer. The general rule is to use half of the total estimated, reduced task times of each feeding path. If the feeding path contains a Critical Chain task, the Critical Chain task is excluded from the calculation because the project buffer already protects it. [Figure 3-5](#) illustrates the placement and size of the feeding buffers for our sample project schedule. The feeding buffer for the upper chain (5 days) is half of the time scheduled for Tasks F and G (10 days). The feeding buffer for the lower chain (7 days) is half of the time scheduled for Tasks H and I (14 days).

[Figure 3-5](#) exhibits two important phenomena unique to CC. Notice first that Task A is not on the Critical Chain, but is a predecessor activity for Task B. Since Task A is a 12-day task, it should have a 6-day feeding buffer. However, that amount of buffer would push the start of Task A to 4 days earlier than the start of the Critical Chain, which is illogical even if possible. Therefore, a dark line in the 6-day feeding buffer denotes the fact that 4 days of the 6-day buffer are *consumed before the project begins*. Some CC scheduling tools add the “days earlier” to the project buffer for additional protection, others simply register the fact that one of the buffers has already been partially consumed, and others push everything out to make room for the buffer. For this example, 4 days have been added to the project buffer, increasing it from 26 to 30 days.

A second item to note is the apparent violation of the practice of starting all tasks as late as possible. In this case, the PM has decided that because Resource 3 on Task I has the possibility of delaying the start of Task C on the CC if Task H and I are delayed more than a total of six days (a distinct possibility since the feeding buffer is seven days), that the lower path in [Fig. 3-5](#) should begin as soon as possible.¹² This action results in a large gap between Task I and the feeding buffer, at the end of which the lower path joins the CC. It is not uncommon for such gaps to occur, given reasoned analysis of risk and additional resource leveling due to the insertion of feeding buffers. Gaps on non-critical paths such as the gaps between Tasks E and F on the upper path in [Fig. 3-5](#) also are no cause for concern.¹³

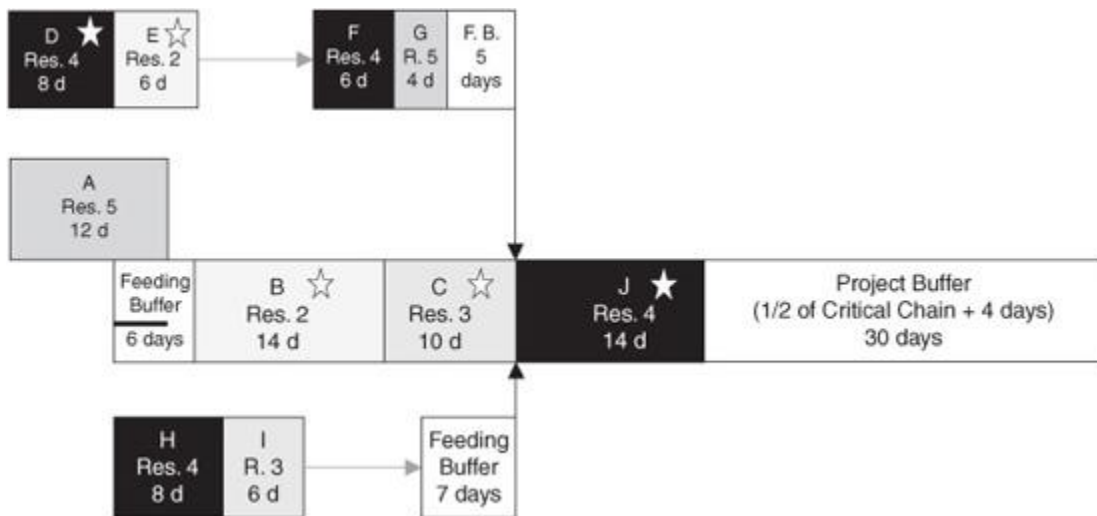


FIGURE 3-5 Critical Chain project schedule with project and feeding buffers.

In terms of the original 5FS, dealing with merging paths would be the subordination step, Step 3.

Another Look at Resource Contention

In order to develop a project plan that has any chance of on-time completion, we must schedule tasks in such a way that the assigned resource is not scheduled to work on more than one task at a time. In CC scheduling, we typically start tasks as late as possible and, when scheduling manually, schedule shorter tasks toward the end of the project when possible. This usually will result in less resource contention as the rescheduling proceeds and provide better opportunities for time recovery earlier during project execution.

As mentioned previously, the critical path in traditional projects may change many times. In CC scheduling, resolving resource contention is doubly important and the possibility of resource contention must be checked in every step of the process.

Looking at the intermediate project schedules in [Fig. 3-4](#) and [Fig. 3-5](#), we see that Task F and Task G are forced earlier in time by the insertion of a 5-day feeding buffer. However, no new resource contention arises due to the insertion of this buffer. Task I, Resource 3, which was pushed earlier by previous action by the PM, is not affected by the insertion of a 7-day feeding buffer. If work on Task I has not been completed by the time Task B (that precedes Task C) is completed, normally the PM will inform Resource 3 to cease work on Task I and move to Task C on the CC.¹⁴ Since Task D is on the CC, Resource 4 first will complete that task, then begin Task H. Should Task D require more than 8 days to complete, Task H might be delayed starting, but the feeding buffer and project buffer can absorb any delays. This simple project example is unusual in that new resource contention does not result from the insertion of feeding buffers. You always should expect new resource contention arising when feeding buffers are added to a project schedule.

Scheduling a resource to work on more than one task at a time can easily result in the resource multitasking in order

to show “progress” on all assigned tasks. Making sure this does not occur in a single project, by leveling all resources, avoids this type of unproductive multitasking. Of course, in a multi-project environment, it is impossible to level all resources over all projects with any confidence that resource contention can be avoided. We must use another CC technique, discussed later, to avoid resource contention in a multi-project environment.

Communications—Step 6

It is imperative that a resource that is assigned to a task on the Critical Chain immediately begin that task as soon as the preceding task is completed. CC uses a notification system that informs the next resource that she or he will be required to work on a CC task. This notification is given a brief time interval before the previous CC task has been completed. In the sample project, this time interval would be two or three days at most.

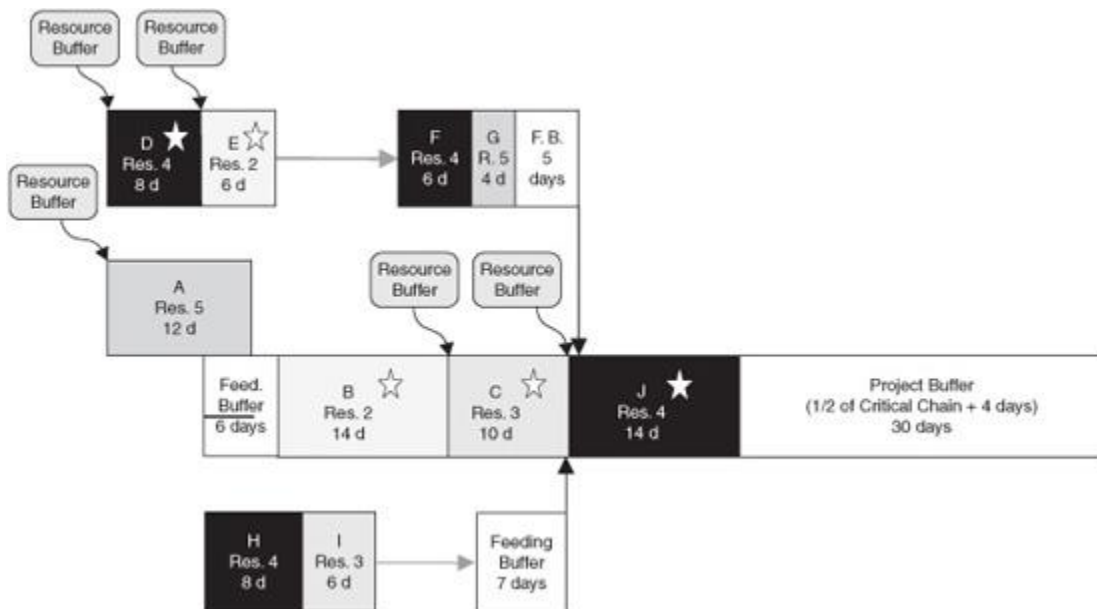


FIGURE 3-6 A complete and fully protected Critical Chain project schedule.

Step 6 of CC project scheduling¹⁵ ensures this notification occurs by placing resource buffers in the project schedule at appropriate points. *Resource buffers* do not have any task time: they *are communication tools*. In addition, resource buffers should be placed in the project plan to inform resources assigned to tasks with no predecessor when they should begin work. Tasks A and D have no predecessors and therefore require early warning signals.¹⁶

The problem of ineffective multitasking was discussed previously. A general policy should be established that states that once a task is begun, it should be completed before another queued task is begun. Certain exceptions can be allowed, such as when the resource must wait for some requirement before he or she can complete the current task. However, the most important exception is when the resource is required on a CC task. The notification time, mentioned previously, should be set at a sufficient time for the resource to “set down” his or her current work in an orderly fashion and prepare for the CC task.

Now we have a fully protected CC project schedule, shown in [Fig. 3-6](#), with no resource contention and with three feeding buffers and a project buffer. The project is now scheduled to complete in 82 days.

There are alternative CC project schedules that are possible for the sample project used in this chapter. This is because the scheduler or scheduling tool may opt to move different tasks forward or backward and thus achieve a somewhat different schedule.¹⁷ The most important concern is not that the schedule is the shortest possible schedule (as most academic literature suggests), but that the promised project completion date is adequately protected.

In [Fig. 3-6](#), *resource buffers* (one or two days) have been placed in the project schedule to notify Resources 4 and 5 when they should begin work on this project. Resource 4 is

informed to begin Task D, then go immediately to Task H. Proper notification (a resource buffer) is given to Resource 2 when work is scheduled to start on Task E on the Critical Chain. Resource 2 is instructed to proceed from Task E as soon as that work is completed to Task B. Like Task H, whose start was transmitted in Task D's resource buffer, a separate resource buffer for Task B is not required.

Even though Resource 3 may still be working on Task I (late completion) when Task B is nearing completion, the resource buffer or other communication about an upcoming CC task may advise Resource 3 to start setting down work on Task I in an orderly way and be ready to begin work on Task C as soon as Task B is completed. Once Task C has been completed, Resource 3 immediately can return to Task I and complete that work. [18](#)

Three Sources of Critical Chain Project Protection

The previous discussion and [Fig. 3-6](#) illustrate that there are three types of protection to improve the likelihood of completing CC projects on schedule:

1. One project buffer of time that can be used for Critical Chain tasks that are not completed in their shortened duration times.
2. Multiple feeding buffers of time that can be used to protect the Critical Chain activities' starts if there are problems with activities on merging paths.
3. Multiple resource buffers that do not add time to the project schedule but provide early warnings to certain resources either to start a path or that they must move to a CC task when needed and sometimes deviate from the standard policy (of not stopping work

on a task until it is completed) in order to start a CC task on time.

In order to present the principles of CC project scheduling, this section considered a simple schedule in a single project environment. We have also presented some clues about basic behavioral changes that are required to make CC project scheduling more effective. Responsibilities for behavioral change will be covered later, but first we will look at the complicated world of scheduling in many or perhaps most environments where many projects coexist.

Scheduling Projects in Multi-Project Environments

A major problem in a multi-project environment is establishing priorities. Not every project can be “Number One.” Setting priorities for projects in a multi-project environment is difficult, but essential. In our experience, many organizations forgo this politically sensitive task and simply cram as many projects as possible into the system in order to take advantage of new business opportunities. Doing so, however, frequently jeopardizes progress on the projects currently underway. The assumption that an early start makes possible an early finish is incorrect. As described previously and in [Chapter 2](#),¹⁹ flooding the organization with projects creates chaos in the project management process, stresses conscientious workers, and tends to burnout the organization’s best people.

Because multitasking is rampant in multi-project environments, and generally is highly valued by management, we wish to stress again its negative effect on productivity. So that you can experience the harmful effects of multitasking, we have included a “Wafer Experiment,” located at www.mhprofessional.com/TOCHandbook that you should conduct. The experiment compares traditional multitasking on three projects with the CC approach. This is

a nice experiment to perform with your children, who may be far better than you may be at manipulating objects on a computer, and who will benefit from being involved in the experiment.

Establishing Project Priorities

It is beyond the scope of this chapter to solve all the problems of prioritization, but it is imperative for every organization in a multi-project environment to use some priority scheme. It does not make sense to permit, by default, the setting of priorities by a resource manager or other person who may not have a global perspective of the organization's many ongoing projects.

Many organizations have established a *Project Management Office (PMO) for the management of their project portfolio*. Some of the possible functions of a PMO are described in [Fig. 3-7](#). Notice the establishment of project priorities based on business priorities, resources, and organizational skills.

Selecting a Scheduling Resource and Establishing Scheduling Buffers

Once project priorities are established, the key TOC concept of buffers can be employed to control the initiation of new projects. In a multi-project environment, each project is scheduled in the same way as in a single project environment, but without regard to resource usage in other projects. Due to massive task duration uncertainty, it is not possible to level all resources across all projects and expect such initial leveling to remain effective for any period of time once project execution is begun.

In order to minimize the need for resources to multitask and to make sure delays on one project do not affect other projects, entry of new projects into the system must be

controlled. We have chosen to use the descriptive terms of “scheduling resource” and “scheduling buffers” in this chapter to restrict the entry of new projects. However, standard terminology has not been established. A search of CC software vendors’ training materials and an investigation of materials and resources used by consultants, academicians, and other CC experts have yielded references to “pipeline buffer,” “staggering buffer,” “drum feeding buffer,” “scheduling resource buffer,” “synchronizing buffer,” “drum buffer,” “sequencing buffer,” “capacity buffer,”²⁰ “drum schedule buffer,” “pacing buffer,” and “capacity constraint buffer.”

Project Management Capability	Business Process Coordination	Project Priorities	Business Metrics
<ul style="list-style-type: none"> • Project Management Practices • Project Management Maturity • Pervasive Project Management Usage 	<ul style="list-style-type: none"> • Strategic View • Goal Alignment • Business Collaboration 	<ul style="list-style-type: none"> • Business Priorities • Resource Application • Leverage Skills 	<ul style="list-style-type: none"> • Progress Reporting • Performance Feedback • Customer Satisfaction

FIGURE3-7 Functions of a Project Management Office. (Reprinted with permission from *A Practical Guide to Earned Value Project Management, 2nd ed., Charles I. Budd and Charlene S. Budd. © 2010 by Management Concepts, Inc. All rights reserved.*)

A *scheduling resource* (SR), somewhat similar to the constraint resource in *Drum-Buffer-Rope* (DBR) implementations for manufacturing, is used to minimize resource conflicts and prevent choking the organization with too many projects. Just as material is scheduled into a

production line based on the system's constraint (the drum that controls the pace of production), we can schedule the initiation of projects into our operations based on the scheduling resource's availability.

Of course, identifying a resource constraint in most multi-project environments is impossible and unnecessary. Therefore, choosing the "right" SR is not critical, but the SR should be one that is utilized across most projects. In software projects, the *integration*²¹ resource typically is chosen as the scheduling resource.

The initiation of each project (in the predetermined priority order) is scheduled such that the SR is leveled across the projects. That is, the SR's²² tasks are never overlapped. New projects can be initiated only at the time that the SR's first task in the new project occurs after the SR's last task in the current project is scheduled to complete.

In addition, we do not want to schedule the SR's tasks in the different projects back-to-back in case one of the tasks should overrun its estimated duration. To provide some protection for the overall multi-project schedule, a *scheduling buffer* is used in each project.

The scheduling buffer is inserted into each project in front of the first task to be performed by the SR. When problems arise in any project, a time buffer in front of the SR's task in the next project will minimize slippage in the entire portfolio schedule. The size of the buffer is optional, but it should be relatively large. Because our entire project portfolio schedule depends on the scheduling buffer, a general rule is to make the buffer at least as large as the duration of recent task times scheduled in the higher-priority project. This is especially true when first establishing a CC multi-project environment. However, buffer size can depend on experience, individual project configurations, and other factors.

For example, suppose we select Resource 4 as a SR. The last two tasks of Resource 4 in the sample project (see [Fig. 3-5](#) or [Fig. 3-6](#)) are scheduled with sequential durations that total 20 days. This project is unusual in that Resource 4 was required to perform four separate tasks on this project. The next priority project only requires Resource 4 to perform two tasks, about the average for this organization. Therefore, the organization has decided that 20 days is a sufficient scheduling buffer to delay the start of Project 2.

[Figure 3-8](#) shows the latter part of the current project, “Project 1,” and two additional projects being initiated as Resource 4 (black color) is available to perform work on them. (Only the latter part of Project 1 and only the beginning part of Project 3 are shown in [Fig. 3-8](#) because the figure is designed to show how the scheduling buffers in Projects 1 and 2 sequence the release of Projects 2 and 3 based on availability of black Resource 4.) The previously described *Strategic-Resource-Buffer* methodology sufficiently staggers the entry of work into the organization’s system, according to a project’s priority, to avoid, to a great extent, any temptation for resources to multi-task. Should there be an occasional example of a resource being required to perform work on different projects at the same time, the PMO or the resource manager can decide which task should have priority.

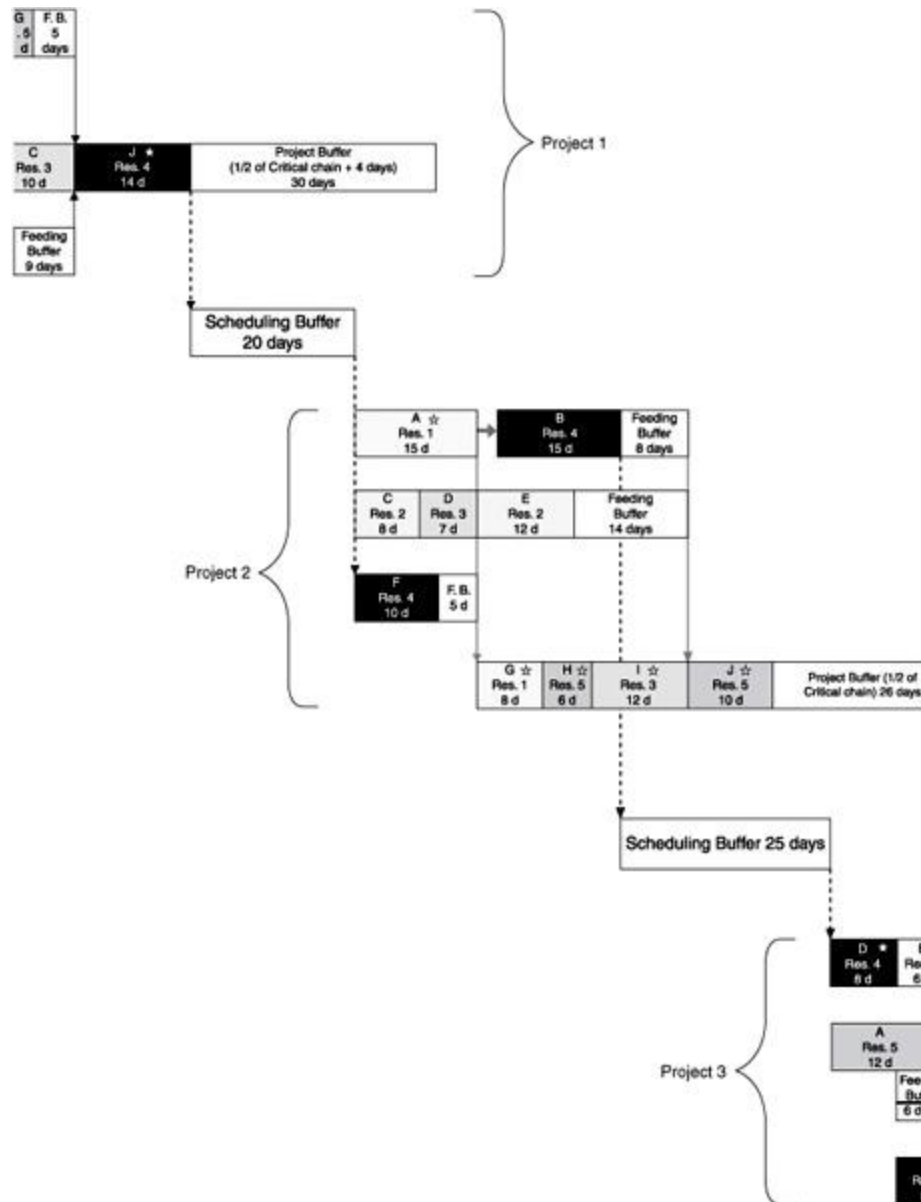


FIGURE 3-8 Scheduling resource (black) and scheduling buffers space the entry of new projects.

The CC in each project in [Fig. 3-8](#) is identified with white stars (*). Project 2 has black Resource 4 scheduled on two tasks for a total of 25 days. Therefore, the scheduling buffer to sufficiently delay the start of Project 3 is 25 days.

Establishing clear project priorities to support an organization's strategy is the responsibility of top management. Priorities should be clear and firm. Should a

more desirable project opportunity arise, project scheduling can be adjusted. However, the impact of delaying projects already scheduled should be computed and considered carefully prior to adding a new project. Change control at the portfolio level is as important as change control on an individual project.

Project Control: The Power of Buffer Management

We previously discussed the purpose of *buffers* as a *project-planning device* to concentrate protection for individual projects and to control the initiation of projects in a multi-project environment. Another very important use of CC buffers is to provide a *project management tool* so a PM knows when to take action and when to avoid doing so unnecessarily.

Tracking Buffer Consumption

To calculate buffer consumption, the PM must have current information on every task that has been started and has not been completed. At each checkpoint (daily or once or twice a week), each project staff member currently working on a task should be asked for the amount of time *remaining* to complete the task. It is unproductive, for project management purposes, to ask for a completion date or percentage of the work that has been completed. (Historically, “percent complete” has often been overestimated.) A “remaining” time estimate is necessary for the PM to know if action is warranted. The remaining time, added to the time elapsed since the task was started, can be compared to the original estimated aggressive time to determine the buffer penetration or recovery. The

reported remaining time *changes* (i.e. is not always decreasing) each time a query is made.

A task duration *overage*, meaning *the task will complete some time beyond the reduced estimated (aggressive) duration*, can be calculated as follows: For a task that has been started and not completed, add the amount of time remaining to complete the task (provided by the assigned resource) to the time elapsed since the task was initiated and compare the current expected total duration to the original estimated aggressive duration. If the current duration is greater than the estimated aggressive duration, the difference between the two is the amount of overage that must be reflected in the appropriate buffer as “used.”²³

The overage calculation is *not* based on when the task was originally planned to start. There is no concern about “start dates” or “finish dates” as each activity time is calculated only on its own planned duration. More will be discussed on this matter later, but as we have already intimated in the preceding section “Communication Plan,” start dates are not emphasized. Instead, CC concentrates on task durations and provides notifications of impending work for each resource on the Critical Chain and for each task without a predecessor task. Otherwise, work is performed in the order in which it arrived in a resource’s queue.

If the overage task is on a feeding chain, the amount of the estimated overage is subtracted from the feeding buffer. If, at some point, a feeding buffer becomes fully consumed, then any remaining overage is shown as utilized in the project buffer. For any task on the Critical Chain, the overage must be subtracted from the project buffer. In a multi-project environment, the project management office (or the equivalent function) should track the performance of the SR (see [Fig. 3-8](#)) so that scheduling buffers can be adjusted if the SR indicates a shorter or longer than planned duration for one of its assigned tasks.

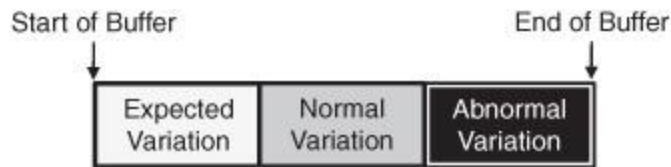


FIGURE 3-9 Buffer variation areas. (Reprinted with permission from *A Practical Guide to Earned Value Project Management, 2nd ed.*, Charles I. Budd and Charlene S. Budd. © 2010 by Management Concepts, Inc. All rights reserved.)

Knowing When to Act

PMs need to have a meaningful knowledge about their project’s status and they need to know when to take corrective actions. The amount of buffer utilization provides the required information. Buffers are generally divided into three equal sections of time that can be thought of as “expected variation,” “normal variation,” and “abnormal variation.” They are somewhat analogous to the green, yellow, and red of a traffic control light. An illustration of this division is shown in [Fig. 3-9](#). In [Fig. 3-6](#), there are 30 days in the project buffer, which means there would be about 10 days in each of the buffer variation sections.²⁴

Expected Variation (Green Zone)

Time has been aggregated in the CC buffers to protect the completion date of the project. If everything works according to the CC schedule, some or all of the buffers will be used and the project will complete on or before the scheduled date. As the project work proceeds, we can expect one-third of the buffers to be utilized due to inherent task uncertainty. That means that, in our sample project from [Fig. 3-6](#), we will expect that 10 or 11 days will be

utilized from our project buffer. No action is required to correct the system at this point. Deming (1993, 194-209) called excessive intervention in operations “tampering.” Taking corrective action when none is required can waste productive time and cause loss of focus.

Normal Variation (Yellow Zone)

The basis for Deming’s discussion of tampering was his hypothesis (now universally accepted) that there are two kinds of variation in any process. He called them “common cause” variation and “special cause” variation (Deming, 1986). Common cause variation is inherent in the design of the process itself because no process is perfect. By their very nature, project task times are uncertain. The utilization of the second third of CC buffers is usually caused due to the inherent uncertainty of task duration prediction. Small variations in the operation of a project are not a reason for alarm, but if the second third of the buffer begins to be used to cover task overages, plans should be formulated to recover lost time. However, to avoid tampering, action should not be initiated until abnormal variation, the last third of the buffer, is experienced.

Utilization of the last (abnormal variation or red) section of a buffer is usually the result of special cause variation and it is wise to observe the scout motto to “be prepared.” The time for the PM to develop an action plan to be used if the red (abnormal) section of the buffer is penetrated is before it happens—while only the second section of the buffer (normal variation) has been penetrated. Among other possibilities, an action plan might include such items as arranging for the possible use of overtime or additional resources, outsourcing parts of the project, or securing an agreement to reduce the scope of the project.

Abnormal Variation (Red Zone)

Special cause (abnormal) variation is usually the result of a unique event outside of the normal course of the project operation. Such events could be as simple as illness of a project resource or as momentous as a natural disaster. When the red portion of the buffer is penetrated, it is definitely time for action and the implementation of the plans made while buffer consumption was in the middle section of the buffer.

If a *feeding buffer* is involved, the appropriate action is to carefully monitor the project buffer. If the project buffer still holds adequate safety, immediate action may not be necessary. If the *project buffer* is involved, the action plan should be initiated immediately. If a *scheduling buffer* is involved, the initiation of the next project should be delayed if possible. Some precedent tasks of the next project may have already been started before knowledge of the SR problem surfaced. If the next project has already been initiated, it would be prudent to delay the initiation of other projects that occur later in the project schedule.

Adjusting Buffers

As a project nears completion, it is expected that some or all of the buffers will be utilized. It is less and less important to maintain the full size of the protective buffers unless they are needed. Remembering that we had to add 4 days from a feeding buffer, the sample project in [Fig. 3-6](#) (or [Fig. 3-5](#)) starts with 30 days in the project buffer. Compared to the original CC time, that is a ratio of $30/52 \approx 0.58$. This ratio of task duration to buffer time should be maintained throughout the performance of the project.

Using [Fig. 3-6](#), for example, when Tasks A, D, and E have been completed, Tasks B, C, and J on the Critical Chain would leave 38 days of work to be completed. Maintaining

the same ratio of 0.58 means that the project buffer can now be adjusted down by 8 days to about 22 (actually 22.04). The traffic light buffer sections of the new buffer would be divided into thirds of 7-1/3 days each and an adjustment made to calculate the new action triggering points. The amount of the reduction in the buffer (8 days) is subtracted from any previous buffer utilization and the difference is applied to the new buffer sections. Assuming that 10 days of safety have been used from the buffer, subtract the 8 days of buffer reduction from the 10 days of safety consumption and the project has utilized 2 days of our recalculated buffer. The project is still “in the green” (experiencing expected variation) and no action is required. See [Fig. 3-10](#) for an illustration of (a) the buffer penetration of 10 days using the original buffer size and (b) the recalculated buffer with two days penetration. (A thick black line denotes the portion of the buffer that has been used.)

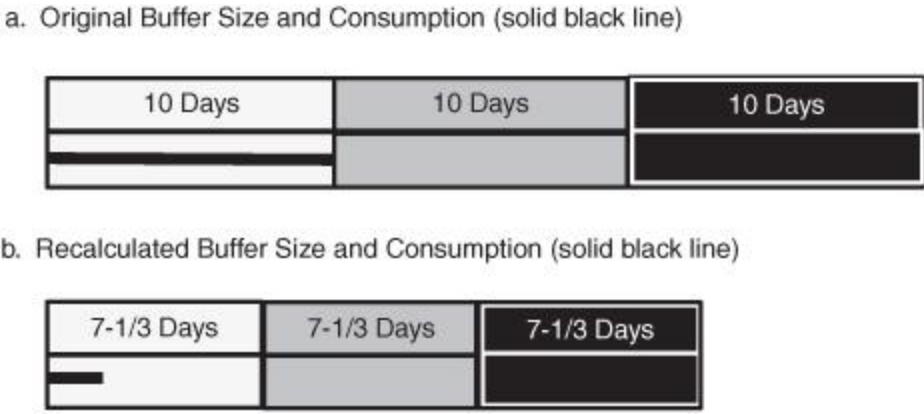


FIGURE 3-10 Original and revised buffer sizes (after completion of Tasks A, D, and E).

Sometimes consumption of the buffer is referred to as *buffer burn rate*. The *TOCICO Dictionary* defines this term as, “The rate at which the project buffer is being consumed in Critical Chain Project Management. The rate is calculated as the ratio of the percent of penetration into the project

buffer and the percent of completion of the Critical Chain" (Sullivan et al., 2007, 7-8). A result of 1.0 would indicate that the original relationship between the Critical Chain and the buffer is being maintained. Using this formula in our sample project, the burn rate would be 0.33 [percentage of buffer consumption: (10 days)/(30 days)], divided by 0.27 [percentage of Critical Chain completed: (14 days)/(52 days)], or 1.22, a bit higher than the desired 1.0. However, [Fig. 3-10](#) indicates the project is still in the expected (green) range of variability.

Feeding buffers similarly are adjusted as the feeding paths are completed. Since Task A has been completed, its feeding buffer is no longer needed. However, 12-day Task A required 16 days for completion, so 2 of the 10 days of the original project buffer used resulted from Task A, which was able to use only 2 days of its original 6 days of feeding buffer.

The PM should know how and why to perform these buffer calculations each time they receive task reports on active tasks, but in complex project settings it would be a very difficult chore without CC project software that reports resized buffers, buffer penetration, and other useful project management information. Various CCPM software programs may compute buffer consumption slightly differently, but the example in [Fig. 3-10](#) will give you an understanding of how the buffers can be adjusted manually as the project is completed.

A typical fever chart showing the trend of buffer consumption versus CC completion over a number of reporting periods, is illustrated in [Fig. 3-11](#). The solid black area (top of [Fig. 3-11](#)) represents the red zone, requiring immediate action, the dark grey area in the middle (diagonal) represents the yellow zone, where plans are made, but action is delayed, and the light grey area represents the green zone where things are going well and the PM should not intervene. Note that by the fourth reporting date, buffer consumption jumped to about 80

percent while only 40 percent of the CC was completed. The project was in the red zone and required immediate intervention. While the project recovered (back to the yellow zone) by the sixth period, additional recovery plans should be formulated in order to ensure on-time completion.

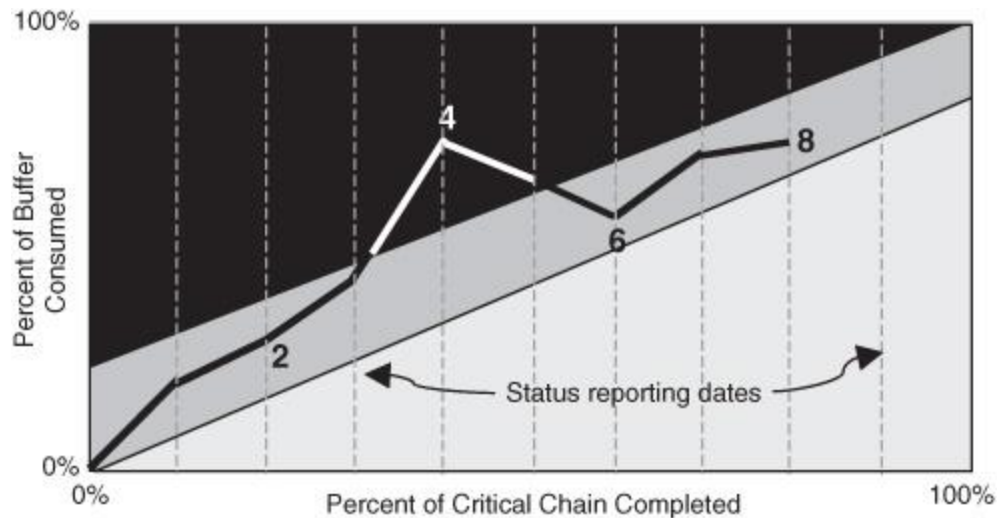


FIGURE 3-11 Buffer tracking on a fever chart [adapted from Newbold (2008), 112].

Using Buffer Consumption Information to Continuously Improve

When buffer consumption enters the *expected* (yellow) variability zone (see Figs. 3-9 and 3-11), every task that overruns its expected (aggressive) time should be analyzed for the cause. This investigation might be initiated for any buffer consumption, starting from the beginning of a project. Causes of overruns (overages) include the following:

- Material damaged or of poor quality
- Resource ill or absent due to family emergency
- Task poorly defined (or poorly understood)
- Quality problem with previous work

- Resource assigned to a more critical project by the PMO (or similar body)
- Subcontractor problems such as poor quality or late delivery
- Unexpected event such as abnormal weather²⁵

Whatever the cause, it should be recorded and all like events aggregated via a check sheet. A Pareto analysis will reveal the most common and expensive causes of delay. This information should be used to analyze how processes and procedures can be changed to avoid future overrun occurrences. The data definitely should not be used to point fingers or berate employees.

Everyone involved in the most common and most critical overrun events should be part of a team to formulate a solution. This group might include individuals involved in predecessor and successor tasks. In this way, an organization can continuously improve its project performance.

Project Budgeting

Now that you have been exposed to CC scheduling and management, we need to return to the subject of project budgeting. We know that we can control task uncertainty with buffers of time. Should we control project costs with a budget buffer of cash? First, let us review, very briefly, a few things about project budgeting.

Keep in mind that the first priority for the organization is completion of every project on or before its CC (shortened) due date. Time is the element that limits organizational profitability. Costs are secondary, or perhaps even further down the list of organization goals. However, if a process is

not established to permit cost savings on projects, they most assuredly will not occur.²⁶

Components of a Project Budget

We are all familiar with the angst of going through the preparation of a regular annual budget and the subsequent budget cycle. Fortunately, preparation of a project budget is much easier and requires fewer schedules. For example, project revenue, either actual or imputed (for internal projects), generally is known prior to a detailed estimation of costs.²⁷ In addition, either the finance or accounting department will take care of managing cash flows, so a project can be treated as a cost center (for internal projects, where only costs are traced to the project) or a profit center (for projects initiated for outside customers and involving revenue generation as well as cost accumulation). Project costs include materials, labor, and overhead.

Materials

Required raw materials, major (costly or unique) supplies, and outsourced work that generally is billed in a lump sum are included in this category and should be estimated for each task that must be performed to complete the project. Materials typically are added when the first task on a path is begun, but can be required for any task.

Equipment purchased for the sole use of the project can be included in the materials or overhead (see subsection "Overhead") category. The original cost of the equipment, minus any resale or salvage value, or, alternatively, the periodic lease cost, should be assigned to the project task designated to use the equipment.

If more than one task requires use of this special equipment, the *net purchase cost* (original purchase price,

less salvage value) or the lease cost of the equipment may be apportioned to the tasks using the equipment by employing a rational and reasonable allocation method. In this case, the project's overhead account is used rather than the materials account.

Labor

Labor can be the largest element of project cost, and includes the *fully loaded* (salary plus benefits) cost of all resources assigned to the project. For convenience, some organizations use an average resource cost per day for all projects, but, with present-day software, it is easy to use individual resource costs. Only the time spent working on the project should be charged to the project.

Overhead

In addition to overhead amounts that may be incurred for the direct benefit of a particular project (such as equipment lease cost or specialized equipment depreciation expense), organizations usually assign a portion of total organization overhead to a project during its life. Overhead costs include information systems, maintenance, and human resources costs, as well as the costs of general materials and equipment commonly used by many projects and departments. Overhead might also include interest expense on borrowed funds.

Good internal control requires direct tracing of all overhead costs to the project benefitting from the use of the overhead item, if possible. However, many organizations use only a few *overhead cost pools* (sometimes called *buckets*) that are basically general ledger accounts where costs of a particular type are aggregated. Then the organization uses simple allocation methods based on

common drivers (allocation bases), such as total materials costs or total labor hours or total labor costs, to allocate overhead amounts. Typically, there are weak cause-effect relationships between the accumulation of costs in the pool, the assumed dependent variable (the item caused by the driver) and changes in the driver (the assumed independent variable) whose increase causes the overhead cost pool to increase.

Typically, overhead costs and driver quantity are estimated prior to the start of an organization's fiscal year for each overhead pool. Then, an overhead rate $[(\text{estimated cost})/(\text{estimated driver quantity})]$ is computed and used to allocate pool costs to "users" of the cost pool. For example, if the estimated annual costs for a particular overhead pool (overhead account) are \$832,000 and the driver is direct labor hours of 208,000, every direct labor hour incurred on a project would be allocated \$4.00 of overhead from this pool.²⁸

A PM should seek to find out everything possible about the organization's overhead allocation process so they can be in a position to negotiate a lower rate if the project does not utilize the services provided by every overhead pool. While difficult to accomplish, some PMs succeed in negotiating a lower overhead rate.

Regardless of how costs are allocated to the project, we now return to the question of how the total project budget should be allocated to project tasks. The remainder of this topic is a bit beyond the beginner level we have assumed thus far, but the following discussion could prove extremely beneficial to your organization.

Assigning Total Project Costs to Project Tasks

Materials naturally are linked to the tasks requiring them and material costs, including outsourced work, which can

easily be traced to particular tasks. Therefore, material costs normally are treated the same for traditional and CC projects.²⁹

Human resource time (labor), however, is another matter. Logically, if aggressive task times are used and resource *time* safety moved to a buffer, costs should follow the same pattern. For example, Task A in [Fig. 3-3](#) required 24 days of work by Resource 5. In a CC schedule, Resource 5 would be asked to complete the task, under different operating policies, of course, in 12 days. Ignoring material costs, if we assume that Resource 5 has a fully loaded cost of \$50 per hour or \$400 per day, \$4800 [(12 days) × (\$400 resource labor cost per day)] would be assigned to Task A and \$2400 (6 days × \$400 per day) would be assigned to the *project budget buffer*,³⁰ analogous to a project buffer of time. While budget buffers might be established for feeding paths as well as for the project, there is little need for such dichotomy between feeding chains (paths) and the Critical Chain when establishing a budget buffer. Therefore, there is a need for only one budget buffer for each project into which half the cost of the safety time removed from all tasks is deposited.

Using traditional product budgeting, Task A would be assigned \$9600 for Resource 5 labor, while CC would assign a total of \$7200 (\$4800 + \$2400) to Task A and the project's budget buffer. The difference in these two amounts (\$9600 and \$7200), or \$2400, would be held at the organizational level as project contingency funds. A PM can freely access funds in the project's budget buffer, but must petition the organization to access project contingency funds. The budget for other tasks would be handled similarly.

For example, the PM could transfer amounts from the project budget buffer to Task A to cover time overages. If Resource 5 required 16 days instead of the estimated 12 days to complete Task A, as we earlier assumed, the CC PM

could transfer \$1600 (4 days × \$400 per day) to cover Task A's overage. Should Task A complete in 12 or less days, the budget buffer would remain intact and remain available to cover other task (or materials) variability.

The project's budget buffer is computed from the safety time and individual resource costs for all tasks on the Critical Chain and feeding paths (chains). If the project completes on or before its calculated due date, any funds remaining in the budget buffer would be returned to Accounting/Finance. If the project requires access to contingency funds, the PMO, or similar body, should provide them upon reasonable and logical petition unless the project is to be cancelled or delayed.

Implementing a New Project Budgeting Process

Quite obviously, for all projects completed for outside parties, an organization would not want to return revenues earned by the completion of a project by its CC due date, which can be significantly shorter than a similar project completed using traditional project management techniques. Therefore, careful attention must be paid to agreed contracts. Contracts not only should state that all revenues promised are earned upon successful completion of the project, but there also may be opportunities to earn a bonus for early project delivery. Likewise, it is not too risky to accept contractual penalties for delivery beyond the promised due date. Even better, these terms (bonus for early completion, penalties for late completion) should be suggested to organizations preparing *Requests for Proposals* (RFPs) so all companies responding face the same terms.

Prior to implementing a new system for allocating project costs to tasks and the establishment of project budget (cost) buffers, however, the CCPM must be implemented and working as expected. Then the 10-step change process

described later and illustrated in [Fig. 3-12](#), shown later, should be followed to ensure that potential negative unintended consequences of such a change do not occur. For example, funds from traditionally budgeted project work may be used by a resource manager to compensate for unbudgeted items such as employee recruitment or specialized training needs, and this situation should be addressed prior to implementing a new budgeting process.

Project Reporting

Communication was one of the *key elements* of CCPM listed early in this chapter. Parts of the internal communication plan were discussed briefly. Communicating (reporting) both internally and externally is presented in this section.

Internal Reporting

The two main aspects of internal reporting are communications among the project team members and information provided to those within the project control functions of the organization. Necessary information is generated from CC Buffer Management. Some aspects of this topic, such as the resource and scheduling buffers, were introduced in the section “Project Control.” Excessive consumption of buffers always should be reported.

Review of Project and Feeding Buffers

The most relevant information to provide inside the company is buffer consumption for all active projects. Consumption of a project buffer that does not match completion of CC tasks (the buffer burn rate) is most revealing.³¹ Because the project buffer serves as an

additional protection for feeding buffers, the ratio of consumption compared to completion of noncritical tasks is not terribly important.

Buffer consumption over time, as revealed in “fever” charts showing which region (zone) of a buffer [green (represented as light grey in chapter figures), yellow (dark grey), or red (black)] has been penetrated at a particular status date, displays trends and charts the history of project operations. As demonstrated in [Fig. 3-11](#), the danger zone (black area) requiring immediate action typically is wide at the beginning of a project and narrow as the project is completed to reflect the decreasing size as work is completed. The pattern of zone sizes over the life of a project, however, tends to vary by industry.

Review of Resource Buffers

A *resource buffer* adds no time to the project plan, but it is a vital part of the CC internal communications plan. It is a calendar event set prior to a Critical Chain task or path initiation so that the PM can inform resources when they should prepare to perform the task. To maintain a tightly planned project schedule, a Critical Chain task must be started as soon as its predecessor is completed. The duration of each resource buffer should be set based on the PM’s knowledge of the project resources and their requirements. Some project members will have comparatively inflexible line duties and will require more set-down and set-up time. All resource buffers, therefore, will not necessarily have the same alert time. Resource buffers are established and placed in the project schedule after the Critical Chain is determined and project and feeding buffers are in place.

Review of Scheduling Buffers

Like project and feeding buffers, a *scheduling buffer* has no task time, but logically is placed between projects to regulate the initiation of projects in a regulated manner. The buffer is used as an internal communications tool to prevent project overload and to limit ineffective multitasking. A SR is designated and scheduling buffers are placed between the SR tasks which staggers the different projects. The buffer size is determined by the PMO (or equivalent) and based on knowledge of the entire project portfolio schedule. If the SR does not have a task in a particular project, the other resources on the project should be examined for potential multitasking and a scheduling buffer inserted before the project is initiated.

Changing Priorities

Periodically, organization priorities will change to admit new projects and to de-emphasize, delay, or cancel projects. Resource managers and PMs need this information so they can establish work schedules.

Since resource managers and PMs typically have access to detailed project information, they may be assigned the task of preparing reports describing the consequences (both expected additional costs and opportunity costs) of changing priorities. Thus, the advantages and disadvantages of a change in priorities can be analyzed prior to a final decision.

External Reporting

External entities in the context of this section will include those that have no immediate and direct oversight of projects and include:

1. The board of directors, the executive committee, or equivalent governing body
2. Other organizations for whom project work has been contracted
3. Regulatory bodies

There are three general requirements for project status reports to external entities:

1. Contractual requirements such as the *Earned Value Management System* (EVS)
2. Non-contractual performance reports to project owners
3. Internal control schedule and scope reports to a project office or similar oversight entity

Project Planning, Deliverables, and Periodic Reporting

In general, the best information to provide in status reports is the buffer burn rate: the percentage of CC tasks that have been completed and the related percentage of consumption of the project buffer. If 25 percent of the Critical Chain has been completed but 40 percent of the buffer consumed, the PM most likely will be required to explain how this trend can be reversed and the project brought back under control.

The Earned Value System

The Earned Value System (EVS) is a set of 32 criteria established for control and reporting in project management. The criteria are organized into five sections with five in Organization, ten in Planning and Budgeting, six in Accounting Considerations, six in Analysis and

Management Reports, and five in Revisions and Data Maintenance. Most branches of the U.S. Government (and other governments) require adherence to the criteria for major project contracts.³² The requiring entities have also published volumes on interpretation of the criteria and “guides” for their implementation. If PMs wish to use the power of CCPM, but are also required to adhere to the EVS criteria and guidelines, they may need to reconcile differences in the two concepts by providing a CC view of status and an EVS view (based on traditional scheduling). Because of the complexity involved, detailed treatment of EVS is beyond the scope of this chapter. However, it is important to note that it is possible to use CC with EVS.

Non-Contractual Performance Reports to Project Owners

If the organization is using only certain aspects of earned value, the use of CC concepts should not present a problem. Buffer Management can be used internally and *earned value* (EV) calculations can be used for external reporting. In a few cases, traditional project plans adhering to EVS criteria have been successfully managed using only the behavioral concepts required for CC (discussed later in the section “Causing the Change”).

Internal Control Schedule and Scope Reports to a Project Office or Similar Body

Legal requirements, such as the Sarbanes-Oxley Act of 2002 (SOX; Sarbanes and Oxley, 2002), have reinforced the need for more stringent internal reporting on projects that impact an organization’s financial reporting system—as do many, if not most, projects. External auditors of companies required to file financial information with the Securities and Exchange Commission now want detailed information on projects that affect the financial reporting process. In our experience, EV,

which is largely based on the familiar standard cost system, is being used more frequently to satisfy these requirements. External auditors are familiar with EV; they may require some education to feel comfortable with CC metrics.

The completion of some projects might determine the life or death of the organization. Other projects have a profound influence on the company's future success. If this is the case, stringent internal controls dictated by SOX (Sarbanes and Oxley, 2002) might apply. Even if a project currently is not required to meet the law's financial operations internal control criteria, it might be required to do so in the future. EVS offers an internal control environment that will meet SOX internal control requirements. Of course, provided auditors have been educated in CC concepts, CCPM also supports SOX internal control requirements.

Causing the Change: Behavioral Issues, Management Tactics, and Implementation

To obtain maximum benefit from CCPM, it should be applied to all of an organization's projects. However, as part of an implementation plan, it may be advisable to conduct a pilot project of one or two projects. While CC concepts may be intuitively obvious to many individuals, you should not underestimate the difficulty of smoothly introducing a new planning, scheduling, and control system.

First, top management must actively and continuously support a CC implementation. Top management must make sure that they, and all other managers, have been trained in the need for new behaviors.

Next, most workers have experienced more than one reorganization and usually many "improvement" programs that have not delivered promised outcomes. However, these are the people who must implement a new system;

management cannot do it alone. The last topic in this section addresses the issue of change.

Managerial Actions to Support Critical Chain Project Management

The more employees who understand CC concepts, the easier a CC implementation will become. However, all managers should receive CC training and should agree with and support CC concepts. In addition, certain actions are required of top managers, resource managers, and PMs.

Top Management Responsibilities

To set the proper environment for change, top managers (CEOs and Executive VPs) will:

- Outwardly and continuously support the implementation of CC concepts.
- Help enforce the ban on unproductive multitasking.
- Reinforce FIFO work rules.
- Direct all performance evaluation, positive and negative, to the project team, *not* to individual project team members.
- Resist any temptation to enter additional projects into the system without appropriate planning, impact analysis, and change control procedures.
- Show appreciation for a crisis-free work environment. (That is, top management will not single out individual “heroes,” who may have solved some crisis, for special recognition.)
- Give attention to sustainable ongoing improvement of project management for the enterprise.
- Use CC schedules and reporting mechanisms to evaluate the implementation of organization strategy and determine required changes.

Resource Manager Responsibilities

Resource managers generally have considerable power in a traditional project management system because they have been able to influence, and perhaps control, project priorities. In a CCPM system, resource managers will:

- Be fully educated in CC concepts so they can make appropriate priority decisions according to buffer consumption reports from PMs.
- Outwardly and continuously support the implementation of CC concepts.
- Work closely with a PMO or similar body in establishing project priorities and selecting a SR.
- Help enforce the ban on unproductive multitasking.
- Reinforce FIFO work rules.
- Emphasize fast turnover of work (analogous to relay race transfers) when a task is completed.
- Enforce the policy of not stopping work once started until it is finished, unless workers receive orders from management (change of priorities) or project status reports indicate that they should work elsewhere.
- Include team performance on projects assigned to individual resources in their overall evaluation.

Project Manager Responsibilities

As front-line managers, PMs should be both capable and creative. PMs will:

- Be available to help any resource that needs help.

- Carefully track all active tasks and immediately record all buffer quantity changes.
- Provide appropriate notice to resources required for upcoming work on a Critical Chain or required to start the first task on a non-critical path.
- Resist the impulse to interfere with the work on a task while buffer consumption is in the “expected” (first third, or green) or “normal” (second third, or yellow) portions of a buffer.
- Formulate an action plan to reverse an unfavorable trend in buffer consumption prior to entering the last third of a project’s buffer.
- Implement planned actions immediately when the remaining buffer is one-third of its expected size according to remaining CC tasks.
- Respect the project priority sequence established by the organization and assist other PMs when possible.
- Enforce the discipline required to protect the project staff from unnecessary multitasking interruptions.

Importance of Trust

Trust is earned slowly and lost quickly. You cannot expect workers who distrust management to welcome any change. A change to CC concepts may be especially difficult if many project workers have considerable experience with traditional project management. With traditional project management systems, tasks appear to require all the time they have been allotted. Estimated task times are a self-confirming prophesy. When people have been scrambling to meet many deadlines, multitasking like crazy, and you tell them they are now working under a new system that requires even shorter durations, they quite naturally will be concerned, if not alarmed. A full explanation of the

anticipated implementation plan, including environmental and other policy changes, is required.

The next topic presents an organizational system to implement change in a way that addresses employee concerns.

Implementing a Critical Chain Project Management System

There is always resistance to change—sometimes for very good reasons. TOC proponents have developed six “layers of resistance” to change (for example, see Kendall, 2005, [Chapter 11](#); Goldratt, [Chapter 20](#), this volume), a familiar topic in behavioral psychology and many other circles. Based on the TOC six layers of resistance, previous behavioral research, and the Budd Innovation Empowerment Model (Budd and Budd, 2010), [Fig. 3-12](#) shows a 10-step process for incorporating concerns and suggestions from many individuals in the organization. [33](#)

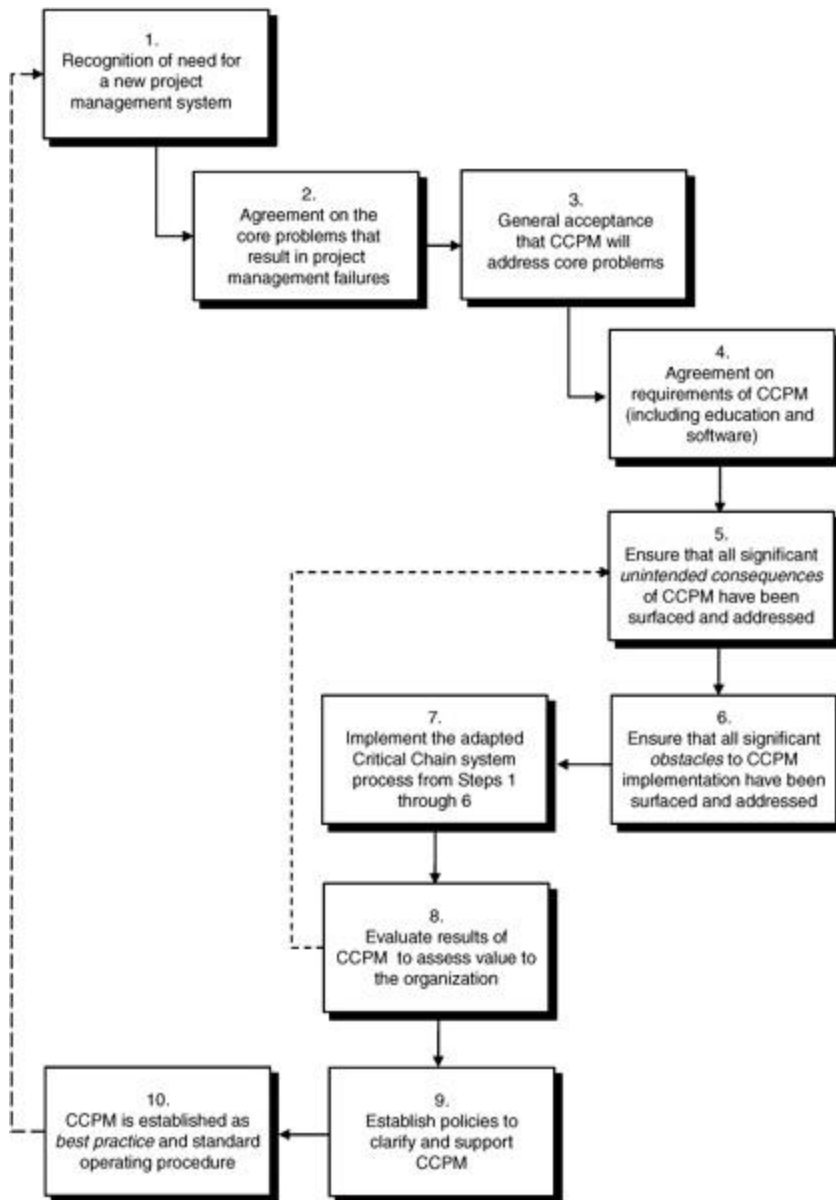


FIGURE 3-12 Critical Chain implementation empowerment model. (Adapted from Budd & Budd, 2010, 260.)

Step 1 in [Fig. 3-12](#) establishes the motivation for change (Why change?). A critical mass of individuals must recognize the pain resulting from continued use of the current system—in this case, traditional project management. Step 2 is the TOC first layer of resistance.³⁴ The remaining steps proceed in numerical sequence. All of the steps must be addressed and none skipped. Some have to be visited more than once

if some members have been left behind at another step or now question a previous step.

The dotted line from Step 8, “Evaluate results of CCPM to assess value to the organization,” to Step 5, “Ensure that all significant *unintended consequences* of CCPM have been surfaced and addressed,” indicates that Steps 5, 6, 7, and 8 may have to be repeated multiple times as implementation proceeds and negative unintended consequences are experienced and overcome.

Once all 10 steps have been taken, a CCPM system is in place and, if no steps have been passed too quickly, the system is working and benefitting the organization as planned. However, as the environment changes, new practices may develop that require changes in the installed system. Therefore, a dotted line also extends from Step 10, “CCPM is established as *best practice* and standard operating procedure,” back to Step 1, signifying the need for a significantly revised project management system. Of course, new improvements in CCPM are being developed every day (see the next three chapters in this book), and your system should be revised from time to time, which may require only a portion of the 10 steps.

Summary

This chapter presents a basic approach to CCPM concepts. Because task times have skewed distributions and cannot be predicted accurately, CC is designed to avoid the dysfunctional behaviors of ineffective multitasking, the student syndrome, sandbagging, and the impact of Parkinson’s Law typical in traditional project management.

To shift concentration from local optima to global optima, safety time is removed from individual tasks and used to protect the entire project. Resource contention is addressed early in the CC planning process and time buffers are used

to address task time uncertainty. Communication tools called resource buffers add to the project communication process and scheduling buffers control the initiation of new projects into the multi-project mix. Full kitting is completed prior to the release of a project. The chapter describes the six regular steps and one optional step in scheduling a single CC project. The three primary sources of safety for on-time project completion are the project buffer, multiple feeding buffers, and multiple resource buffers.

In multi-project environments, it is essential to have a prioritizing process for projects. None of the projects will complete on time if there are so many projects that resource scheduling is difficult and multitasking is rampant. A “Wafer Experiment” located at www.mhprofessional.com/TOCHandbook is an excellent way to experience the effects of bad multitasking. A SR is similar to the constraint resource in DBR implementations in manufacturing. A scheduling buffer, based on the SR’s availability, will minimize resource conflicts and prevent choking the organization with too many projects.

Buffer Management gives the PM important information on the project status. When actual task durations are longer than planned in the project schedule, the overages are subtracted from the buffer. Normal variations in task durations are expected to consume some or all of the buffer time during the operation of the project. An extreme rate of buffer consumption can inform the PM when extraordinary action is required. As the project is completing, the size of the buffers can be reduced as there will be less and less task protection time required.

The use of time buffers in the project schedule has been covered extensively and the use of budget buffers in planning the project might be helpful. Typically, project budgets are derived from the costs of materials, labor, and overhead. As components of the project schedule are moved to time buffers, the associated costs could be moved

to budget buffers. Careful attention must be paid to contract wording so that unintended consequences do not occur because of early (or late) project completion.

Internal reporting in CCPM is accomplished primarily with buffer reports. For external reporting, either CC metrics or a formal or informal EVS may be used.

Implementing CCPM will require changes in the typical behavior of project team members and in organizational policies and procedures. Certainly, management support is crucial and a pilot program might be advisable. The chapter describes the responsibilities for top management, resource managers, and PMs and reinforces the need for intraorganizational trust. Because there will always be some resistance to any change, a CC implementation empowerment model graphically illustrates the steps in overcoming resistance and dealing with unintended consequences.

References

Atallah, M. J. 1999. *Algorithms and Theory of Computation Handbook*. Boca Raton, FL: CRC Press.

Budd, C. I. and Budd, C. S. 2010. *A Practical Guide to Earned Value Project Management*. 2nd ed. Vienna, VA: Management Concepts.

Defense Contract Management Agency. 2009. *Earned Value Management Systems Criteria*. In Defense Contract Management Agency [database online]. Available online at <http://guidebook.dcma.mil/79/criteria.htm>.

- Deming, W. E. 1986. *Out of the Crisis*. Cambridge, MA: MIT Center for Advanced Engineering Study.
- Deming, W. E. 1993. *The New Economics for Industry, Government, Education*. Cambridge, MA: MIT Press.
- Goldratt, E. M. 1990. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Crotonon-Hudson, NY: North River Press.
- Goldratt, E. M. 1997. *Critical Chain*. Great Barrington, MA: North River Press.
- Goldratt, E. M. and Cox, J. 1984. *The Goal: A Process of Ongoing Improvement*. Great Barrington, MA: North River Press.
- Kendall, G. I. 2005. *Viable Vision: Transforming Total Sales into Net Profits*. Ft. Lauderdale, FL: J. Ross Publishing.
- Leach, L. P. 2005. *Critical Chain Project Management*. 2nd ed. Norwood, MA: Artech House.
- NASA. 2009. *NASA Schedule Management Handbook Draft*. (Revision 16a, April 3). Online: National Aeronautics and Space Administration.
- Newbold, R. C. 2008. *The Billion Dollar Solution: Secrets of Prochain Project Management*. Lake Ridge, VA: ProChain Press.
- Parkinson, C. N. 1957. *Parkinson's Law*. Boston: Houghton Mifflin.
- Rubinstein, J. S., Meyer, D. E., and Evans, J. E. 2001. "Executive control of cognitive processes in task

switching.” *Journal of Experimental Psychology: Human Perception and Performance* 7(4) (August):763–797.

Sarbanes, P. S. and Oxley, M. G. 2002 . *Sarbanes-Oxley Act of 2002, H.R. 3763*. Washington, DC.

Shellenbarger, S. 2003. “Multitasking makes you stupid: Studies show pitfalls of doing too much at once.” *Wall Street Journal*, February 27, sec D.

Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>

United States Government Accountability Office. 2009. *GAO Cost Estimating and Assessment Guide: Best Practices for Developing and Managing Capital Program Costs*. Washington, DC: GAO-09-3SP.

About the Authors

Charlene Spoede Budd is a Professor Emeritus from Baylor University, having taught management accounting and project management classes for a number of years. She received her undergraduate degree (accounting major, *Summa Cum Laude*), and MBA degree from Baylor University (1972 and 1973, respectively), and her PhD from The University of Texas at Austin (1982), where she specialized in the fields of accounting, economics, and finance. She holds the following active professional designations: CPA, CMA, CFM, PMP. In addition, she is certified in all areas of the Theory of Constraints by the Theory of Constraints International Certification Organization (TOCICO).

Her research has been published primarily in practitioner journals and she has been awarded three Certificates of Merit for articles published in *Strategic Finance*. She also has singly or coauthored publications in *Industrial Marketing Management* (special issue on projects), *Human Systems Management Journal*, *Today's CPA*, *The Counselor*, and other journals, and many conference proceedings. Charlene has coauthored two accounting textbooks and she and current coauthor, Charles Budd, have published *A Practical Guide to Earned Value Project Management*, 2nd Edition (Management Concepts, 2010) and *Internal Control and Improvement Initiatives* (BNA, 2007).

She is active in several professional organizations, including the American Accounting Association, Financial Executives Institute, and the Project Management Institute. In addition, she has been a member of the AICPA's Content Committee and was chair of the Business Environment and Content Subcommittee of the AICPA for the past several years. Currently, she is chair of the Finance and Metrics Committee of the TOCICO.

Most of her time now is devoted to research, but she also is a member of the Board of Directors of a public company.

Dr. Janice Cervený is on the faculty of the College of Business, Department of Management Programs. She has worked primarily in the blood banking and health care industries but now consults and trains many diverse organizations in the Theory of Constraints. She is an Avraham Y. Goldratt Institute "Jonah," "Jonahs Jonah," and is certified in the functional-specific applications of TOC for Production (Drum-Buffer-Rope, DBR), Distribution/Supply Chain Management (Continuous Replenishment, CR), Project Management (Critical Chain, CCPM), and interpersonal management skills applications (Management Skills Workshop, MSW).

She has had a number of for profit and not-for-profit clients in the South Florida area including NCCI (National Council on Compensation Insurers), Siemens Telecom Networks, Sensormatic Electronics Corporation, Office Depot, the North Broward Hospital District, and Philips Electronics. She has most recently completed a contract with the Veterans Administration in Washington, DC, for clinical practice managers resulting in her editing a book for Ambulatory Care Clinic Managers. Her most recent article (with Dr. Stuart Galup) "Critical Chain Project Management: Holistic Solution Aligning Quantitative and Qualitative Project Management Methods" appeared in *Production and Inventory Management* [43(3&4):55-64, 2002].

She is a member of the American Production and Inventory Control Society (APICS), the Decision Sciences Institute (DSI), the American Society for Quality (ASQ), and the Theory of Constraints International Certification Organization (TOCICO). She is recognized by the latter as internationally certified to facilitate implementations of TOC applications and is chairman of the TOCICO Project Management Certification Committee.

She received her undergraduate degree from the University of Texas at Austin. Her PhD is from the State University of New York at Buffalo's School of Management.

CHAPTER 4

Getting Durable Results with Critical Chain—A Field Report

Realization Technologies, Inc.

Background

“Overdue and over budget” is what most often comes to mind whenever one mentions “projects.” An equally depressing image is one of long hours, firefighting, and chaos. It is against this backdrop that Critical Chain was introduced by Dr. Eliyahu Goldratt in 1997.

Since 1997, Critical Chain has been deployed in a wide range of organizations. Many of them have achieved results that are nothing short of amazing—whether they are in the private sector or public; engaged in blue sky R&D or industrial projects; large or small; or based in western or eastern countries. Some of them have won top honors including the 2006 Franz Edelman Award,^{[1](#)} and the 2009 TOC North American Achievement Award.^{[2](#)}

Purpose and Organization

Like all improvements, the concepts of Critical Chain are straightforward. However, just like other breakthroughs, the science behind a concept doesn't automatically deliver its benefits. The results shown in [Table 4-1](#) came from “engineering” the nitty-gritty details of putting Critical Chain

concepts into practice. To paraphrase, success was 1 percent science and 99 percent engineering.

Copyright © 2010 Realization Technologies, Inc.

Project Descriptions	Before	After
Design, development, and upgrade of telecommunications switches (300–400 active projects, 30+ deliveries a month)	Lead times were long On-time delivery was poor 2000 people	10 to 25% reduction in lead times 90+% on-time delivery 45% increase in productivity
Design and manufacturing of oil and gas platforms	Design engineering: 15 mos. Production engineering: 9 mos. Fabrication and assembly: 8 mos.	Design engineering: 9 mos. Production engineering: 5 mos. Fabrication and assembly: 5 mos. 22% higher labor productivity
Pharmaceutical research and development	Completed 5 projects per quarter in 2005 55% projects delivered on time	12 projects per quarter in 2008 90% projects on time No increase in resources
Customized customer billing and management systems for the telecommunications industry	Market pressure to reduce project cost and cycle time	Increased revenue per person for 4000 people by 14%, reduced project cycle time by 20%
Steel plant maintenance	Boiler conversion: 300–500 days Routine maintenance and upgrade took too long	Boiler conversion: 120–160 days Reduced maintenance and upgrade durations by 10 to 33% in Year 1, and another 5 to 33% in Year 2

New product development (home appliances)	34 new products per year 74% projects on time	Increased Throughput to 52 new products in Year 1 and 70+ in Year 2, with 88% projects on time and no increase in head count
Helicopter repair and overhaul	H-46 aircraft turnaround time: 225 days H-53 aircraft turnaround time: 310 days Throughput: 23 per year	Reduced H-46 turnaround time to 167 days, with more scope Reduced H-53 turnaround time to 180 days Delivered 23 aircraft in 6 mos.
Repair and overhaul of C-5 aircraft (cargo planes used by the U.S. Air Force)	Turnaround time: 240 days 13 aircraft in repair cycle	Turnaround time: 160 days 7 aircraft in repair cycle 75% reduction in defects
Equipment for manufacturing solar panels (engineer-to-order)	Revenues of €130 M Profits of €13 M Cycle time: 17 weeks 80% on-time delivery	Increased revenues to €170 M Increased profits to €22 M Reduced cycle time to 14 weeks 90% on-time delivery

Source: Presentations by respective organizations at Realization's Project Flow Conferences, 2004–2009 available at www.realization.com/results, where more examples can be found.

TABLE 4-1 Examples of Critical Chain Results

The purpose of this chapter is to share how successful adopters have put Critical Chain concepts into practice and achieved durable results. It is based on experience from more than 200 enterprise-level³ implementations of Critical Chain. The range of these includes development of high-tech products; R&D and commercialization of pharmaceuticals; IT applications; design and manufacturing of complex equipment; shipbuilding; building, erection, and commissioning of physical infrastructure; and maintenance, repair, and overhaul of aircraft, submarines, and ships, as well as steel plants and oil refineries.

Starting with a quick recap of Critical Chain, this chapter discusses practical challenges in implementing it successfully. Then, a step-by-step process of implementation is described, followed by an overview of lessons learned over the last 12 years. Finally, before the summary, there are answers to the frequently asked questions that have not been covered in the rest of the chapter.

Recap of Critical Chain

Executing projects is like conducting an orchestra. Various inputs, resources, equipment, decisions and corrective actions have to be brought together at the right place and the right time throughout the life of a project.

Unfortunately, uncertainties get in the way. Tasks take longer, vendors don't deliver on time, technical glitches happen, requirements change and so on. As these uncertainties unfold, even the most carefully prepared plans go awry. Execution priorities become unclear (which tasks to do first) and unsynchronized (every department, every person starts prioritizing their tasks differently).

Consequently, a project is mostly *waiting* for something or the other (see [Fig. 4-1](#)). For example:

- *Waiting* for resources because they have been assigned to other tasks.
- *Waiting* for specifications, approvals, materials, etc., because the supporting resources that were supposed to supply or obtain them were busy elsewhere.
- *Waiting* for issues to get resolved because experts are firefighting other issues.
- *Waiting* for decisions because managers have too much on their plates.
- *Waiting* for all feeding legs of the project to come together at integration points.



FIGURE 4-1 Time-traps in projects.

As these wait times accumulate, projects become late, firefighting starts, and resources are pulled in multiple directions at once. Priorities keep changing and people are forced to multitask.⁴ Managers' ability to control outcomes is compromised and they often suffer a near-total loss of

control. They cannot predict when a project will finish because holdups keep happening; and they don't know how much capacity is really needed because no matter how many resources they provide, resources are still overloaded while projects continue to run late.

The net impact is that projects take much longer than they should, deliver less scope than originally planned, and are costlier than they need to be. In addition, resources are less productive than they might be.

Critical Chain solves *all* these problems by synchronizing task priorities within and across projects, and within and across departments. To synchronize, Critical Chain uses three precepts, or Rules:

1. **Pipelining:** Limit the number of projects in execution at one time.
2. **Buffering:** Discard local schedules and measurements, and use aggregate buffers⁵ to protect against uncertainties.
3. **Buffer Management:** Use buffer consumption to measure Execution, and to drive execution priorities and managerial interventions.

Rule 1 Pipelining: Limit the Number of Projects in Execution at One Time

When too many projects are in execution compared to available capacity—hereafter referred to as *high work-in-progress* or high WIP—it automatically causes execution priorities to become unsynchronized.

For example, if several projects are simultaneously in execution, different departments might prioritize their work differently. All projects can make some progress but then become stuck at integration points where work-streams

from different departments have to come together. Task priorities within departments could also get unsynchronized, in which case even the departmental work-streams would take longer. Unsynchronized priorities also create schedule conflicts, which can cause the individual resources to multitask, which results in lower quality.

If fewer projects are in execution, the chances are much higher that task priorities within and across departments are synchronized. The higher the WIP, the smaller the chances that task priorities will be synchronized!

Therefore, the first rule for execution success is: *Limit the number of projects being run at a time*. Projects should be staggered based on the most limiting resources because at any time only as many projects can be executed as you can get through those constraints. Any extra projects will only spread resources more thinly and destroy synchronization. Enforce this rule even if it means leaving some resources idle!

Rule 2 Buffering: Discard Local Schedules and Measurements, and Use Aggregate Buffers

The traditional project management approach is to turn task schedules and estimates into commitments. It assumes that if people are held accountable, they will finish individual tasks on time and on budget, and the entire project will consequently be on time and on budget.

Unfortunately, this traditional approach only leads to longer projects while causing execution to become more unsynchronized:

- In planning, accountability for task-times causes people to include contingencies in their commitments—they have to plan for uncertainties as well as the reality that

most of this time will be spent waiting for one thing or another. That is how project plans are extended.

- In execution, resources now not only are scattered across too many projects, but also have an incentive to work on easy tasks—tasks that will help them beat or meet their estimates—instead of working on tasks that are most critical to the project.

Therefore, the second rule for execution success comes down to: *Allow individual tasks to exceed their planning estimates*. To protect projects from task delays, buffers are inserted before integration points and at the end of the project. With lower WIP, with the pressure to meet estimates gone, and with buffers to take care of uncertainties, the contingencies embedded inside task estimates are no longer needed and can be stripped out.

Not only does this second rule allow for shorter project plans (because buffers are smaller than the sum of task-level contingencies), execution becomes easier as well. With shorter project plans there is significantly less pressure to start projects as soon as possible; extra time can be used to get ready for execution through better preparation.

Rule 3 Buffer Management: Use Buffers to Measure Execution, and Drive Execution Priorities and Managerial Interventions

With low WIP and adequately buffered project plans, a single priority system can be firmly established in execution. The essence of the third rule is simple, but profound: *Prioritize tasks according to buffer consumption*. The highest priority is given to project legs that are consuming buffers at the fastest rate.⁶ When every person and department follows these priorities, they are all synchronized—automatically!

Buffer-based priorities not only are synchronized, but they also cause project status to be reliable. If resources work on

the right tasks at the right time, it is assured that current project status is an accurate predictor of the future—despite uncertainties, most of which can be absorbed into the properly sized buffers. If recovery actions are initiated whenever buffers are showing “over consumed,” many abnormal uncertainties can also be combatted.

Practical Challenges in Implementing Critical Chain

Experience shows that no matter what the environment, there are three sets of challenges in realizing the benefits of Critical Chain. These challenges all arise from the fact that Critical Chain is an *enterprise solution* for synchronizing project execution, rather than a planning and control technique for individual project managers.

Challenge 1: Gaining Managerial Commitment for Implementing the Three Rules

To state the obvious, without managers’ commitment it is not possible to activate any new management rules.⁷ To be clear, commitment is not about managers agreeing with the idea of Critical Chain. It is about them thoroughly thinking through the details of the changes, overcoming the hurdles that will come up, and getting results.

Buy-in needed to gain the commitment: As many would attest, even after managers are trained by experts and even after the method has been successfully piloted on one or two projects, organizations may not undertake a full implementation. Not surprisingly, lofty visions and abstract mission statements advocated by change management gurus don’t break the inertia either.

True buy-in is achieved only when managers realize why improving project performance is vital for the business (why change?). They also need to appreciate that the

management challenges they face daily and the inordinate waste of time and capacity stem from the same root cause, that is, poor synchronization of tasks and resources.

Challenge 2: Translating Concepts into Practical Procedures and Instructions

Once managers have bought into the need for change and the validity of the Critical Chain Rules, a host of technical questions arises. What is the right level of WIP? How do you transition from high WIP to low WIP? When do you release new projects into execution? How do you size buffers? How much detail do you put into project plans? How do you ensure that removing local due dates and local efficiency measurements does not lead to loss of accountability? What does it mean to actively manage buffers?

Many such questions are answered throughout this chapter in a summary form. “TOC Insights into Project Management”⁸ and “The Goldratt Webcast Program on Project Management”⁹ provide more in-depth explanations.

Challenge 3: Sustaining the Critical Chain Rules and Results

How are organizations prevented from sliding back into their old mode of running projects? How do you adjust Execution as business needs change? Can the implementation be protected from changes in personnel, especially at the top?

These issues are not unique to Critical Chain but common to all management systems. Moreover, sustained superior performance is not a natural state for organizations; it requires strong leadership to produce great results on an ongoing basis. Still, hands-on experience in many environments has repeatedly shown that the following actions significantly increase the odds of sustained success with Critical Chain.

Mechanizing the Changes

Embedding the Critical Chain Rules into management *policies*, management *processes*, and management *information* makes an implementation less dependent on people. It makes sure that the Rules are not subject to individual choice; and it also allows them to be easily understood and translated into decisions and actions. Following are some examples of such mechanization, and making the fruitful practices routine:

- **WIP Policy:** Set a limit on the number of projects that can be in execution at a time.
- **Pipeline Reviews:** Are WIP limits being followed and, if not, why?
- **WIP Alerts:** Highlight if the actual WIP exceeds the allowed WIP.
- **Task Management Policy:** Make task managers accountable for following the buffer-based task priorities.
- **Task Management Reviews:** Are task managers assigning resources in order of priority and, if not, why?
- **Priority Compliance Reports:** Show if tasks are being worked out of priority.

Establishing a Process of Ongoing Improvement

With Critical Chain, improving performance should not be, and does not have to be, a onetime event. Analyzing buffer consumption highlights the problems to solve to keep getting incremental improvements in overall performance. For example, a leading provider of food packaging increased its Throughput from 72 sales projects per year to 116, and

then from 116 to 171. Originally completed in 2003, this implementation is still going strong.

Identifying improvement opportunities by analyzing buffer history ensures that local improvements will not only have a global impact, but also not violate the Rules. Actually, making those improvements will only increase the value of the Rules.

Turning "Execution" into a Business Asset

Improving performance is not just about catching up with the backlog and due-dates. It is also about building a business advantage. Once project speed, efficiency, and predictability become a business asset (high margins, low investment in operating infrastructure/equipment, or a competitive edge), the pressure to sustain results, as well as the Rules that enable them, will not subside.

For example, when the U.S. Air Force got used to having fewer aircraft in maintenance and more aircraft available to fly missions, its logistics centers *had* to sustain the fast turnaround times—the frequent changes in their military leadership notwithstanding.

Similarly, after a large provider of IT applications to the telecommunications industry got used to 14 percent higher revenue per person and 20 percent shorter project durations, the part of the organization delivering those improved results was not only expected but *required* to continue performing at those levels. Even more impressively, as increasing globalization and the 2009 downturn in the economy put more pressure on prices, this group rose to the challenge once again and was instrumental in maintaining corporate profitability.

Step-By-Step Process for Implementing Critical Chain

This section describes seven practical steps developed in the field over the last 12 years for getting results quickly (in weeks, not months or years).

The operative word here is “quickly,” not only because results can be achieved quickly but because the actual results (or lack thereof) also provide useful feedback to the implementation teams. In those implementations where results were being achieved quickly, it increased confidence and strengthened buy-in; and in cases where anticipated results were not being achieved, course corrections could be made early.

It does not matter whether an organization is large or small; results can be realized within weeks—in fact, as soon as the first rule of low WIP is put into practice.

The seven-step process is as follows:

Step 1: Achieve management buy-in

Step 2: Reduce WIP and implement “full kitting”

Step 3: Build buffered project plans

Step 4: Establish task management

Step 5: Implement surrounding processes

Step 6: Identify opportunities for continuous improvement (POOGI¹⁰)

Step 7: (When applicable) Use superior delivery as a competitive advantage to win more business.

Each step is discussed in more detail in the following sections.

Step 1: Achieve Management Buy-In

Experience confirms that the TOC buy-in process¹¹ works quite well, especially when facilitated by skilled and knowledgeable implementers—people who know the details of the adopter’s business and operations as well as Critical

Chain. For our purposes, this process translates into getting the following five agreements from management:

1. We need to improve project execution.
2. The solution lies in synchronizing resources by implementing the Three Rules.
3. The Three Rules can be translated into practical procedures.
4. We can take care of all the possible negative side-effects (e.g., loss of accountability when task level measurements are discarded; project managers gaming the priorities by manipulating buffers in their projects; delaying discovery of risks by not starting projects as soon as possible; etc.).
5. All the implementation obstacles (e.g., potential conflicts with “earned value reporting”) can and will be addressed.

As the TOC buy-in process of gaining unequivocal acceptance is widely known and documented, this chapter will not dwell on it further.

The major learning to emphasize is that instead of pursuing Critical Chain as a “best practice,” successful adopters used business needs to drive the implementation. Critical Chain was viewed as a means to this end. They analyzed how project performance was linked to making more money or achieving more of the goal of their enterprise, quantified the gap between desired project performance and actual project performance, and set improvement goals accordingly. See [Table 4-2](#).

To demonstrate that managers are committed, the improvement goals they set should be ambitious. It has been consistently observed that organizations are more

easily galvanized toward ambitious goals than around incremental improvements. Moreover, only when ambitious goals are set are substantial improvements realized. Modest targets are rewarded with moderate results, and lack of targets is accompanied by absence of results.

Improvement goals are usually set for higher Throughput (i.e., how many more projects, features, experiments, studies, etc., a year compared to current performance) and for faster cycle time.

Project Types	Examples	Link between Project and Business Performance
Project-Based Businesses	Equipment manufacturers Maintenance, repair, and overhaul (MRO) operators IT services providers Engineering services firms	Generate higher revenues and profits using the same resources. Use delivery performance (deliver faster and on time) to win sales and even charge premiums.
Internal Projects: New Product Development	High tech pharmaceuticals	Get to market fast to capture market share and charge higher price. Segment the markets/play in more segments by offering many products.
Internal Projects: Infrastructure Setup	New factories, stores, bridges, railway lines, etc.	Get to market fast to start generating return on investment faster. Sometimes capture market share and charge higher price.
Internal Projects: Maintenance, Repair, and Upgrade of Captive Assets	Airlines and defense establishments with their own MRO facilities Process manufacturing plants like oil refineries and steel plants	Increase productive uptime by doing maintenance and upgrades faster. Increase productive yield through higher quality of maintenance.

TABLE 4-2 Links between Project and Business Performance for the Basic Types of Projects

In just one real life example from among many, leadership of a large military organization set a target of increasing the number of testing projects by 40 percent—even though project people were overloaded and projects were running behind. Within three months, the organization was delivering 25 percent more projects, with 30 percent

reduction in cycle times. In addition, the goal of 40 percent increase in Throughput was realized in eight months.

Step 2: Reduce WIP and Implement “Full Kitting”

Since the traditional mode of operation leads to too many projects in execution, there are two aspects to pipelining: one is transitioning from high WIP to low WIP, and the second is maintaining low WIP by releasing projects in a metered fashion. This step is about transitioning from high WIP to low WIP.¹²

The typical process for transitioning from high WIP to low WIP is as follows:

1. Create a list of projects in the various phases of execution. These phases in different types of projects, for example, can be:

- IT projects—Scoping, design, coding, and unit testing; system testing; and user testing
- New product development—High-level design, low-level design, virtual testing, prototyping, physical testing, and production ramp up
- Engineer-to-order—System design, detailed design, procurement, manufacturing, and assembly and testing
- MRO—Inspection and disassembly; repair, assembly, and inspection; and trials

2. Specify one of the phases as the *drum*.¹³ Drum is the phase that can accommodate the least number of projects at a time. Put 25 to 50 percent of the workload temporarily on hold in the overall pipeline as well as the drum. There is no need to worry about

selecting a wrong drum at this time (it can be corrected later), or to be exact in calculating workload from each project. The objective is simply to free up enough resources so that remaining projects will be done substantially faster.

3. For the time being, organize remaining projects using a simple priority process like project due dates. The project that is due first gets the first shot at resources; remaining resources are given to the project due next; and so on. This will accelerate the rate of completing projects. A sophisticated process for synchronizing resources (i.e., based on the rate of buffer consumption) is implemented later (see Step 4).
4. Deploy any unassigned resources to “full kitting” the projects on hold. Full kitting is the process of clarifying requirements, getting sign offs, staging of materials, etc. It is important to distinguish between full kitting and actually doing the tasks: activities that allow project tasks to be done without interruptions are included in the full kit list, whereas activities that directly progress the tasks are excluded.
5. As in-process projects are completed, release the on-hold projects one by one according to their established priority.

Avoid paralysis by analysis. The goal is to get results quickly. For example, a major pharmaceutical company almost doubled its rate of project completion in the first 8 weeks—following the implementation of this step, they finished 11 projects compared to 6 projects in the prior 8 weeks. In other cases of comparable potential, hesitancy led to “more study,” initial enthusiasm and momentum were

lost, and the implementations were subsequently abandoned.

Step 3: Build Buffered Project Plans

Project plans are needed to provide execution priorities and early warning signals, and require the following data:

- Tasks and dependencies (precedents and hand-offs) among tasks
- Duration of tasks
- Type and quantity of resources needed for each task
- Task managers
- Buffers (feeding buffers, contractual milestone buffers, and project buffers)
- Resource types and the maximum units of a resource type available to the project
- Project end date and contractual milestone dates

While the concept of a project plan is simple, many organizations struggle with defining the right level of detail.

Degree of Detail Required in the Plans

Too many tasks in a project plan induce multitasking, make it difficult to analyze plans and buffer consumption, and generally lead to loss of control. Not enough detail, on the other hand, leads to unclear priorities and hence the same effects.

Based on evidence from a wide range of industries, a task should be 3 to 7 percent of a buffered project's lead time. More than 250 to 300 tasks in a complex project and less

than 10 to 15 tasks in a simple project are not recommended. If this guideline yields tasks that are too long and thus not useful for task managers, then subprojects can be used to zoom into detailed tasks instead of adding tasks to the main project. Subprojects that are two to four weeks long can be a network of tasks without buffers, while longer subprojects must be properly buffered project plans. For reference purposes:

- An IT application development project with 200 people working on it for four months is executed successfully with 150 tasks.
- Aircraft maintenance and repair projects with 50,000 hours of work per project and durations of 7 months are managed with less than 200 tasks.
- A 3-year pharmaceutical research and development project, with about 50 scientists and professionals working on it, could be managed well with about 175 tasks in the main project.
- Development of digital cameras with 100 engineers and 10 projects at a time could be managed with 150 tasks per project.
- Ten-day long helicopter maintenance and repair projects requiring about 4500 hours of work are being managed with 15 tasks.
- Commercial shipbuilding projects with 750 people working on a project for one year are being managed with 275 tasks.
- Construction of five 50-story buildings with 1600 people working on them for three years was managed with 5 projects and 290 tasks in each project.

1. Define cycle-time targets.
2. Communicate to all managers that people will not be measured in execution against the task estimates used in planning.
3. Assemble a team of representative project managers and task managers and conduct a workshop to get their buy-in into the possible gains with the Three Rules.
4. Create project plans without buffers.¹⁴
 - Define the project's objective and the task that will signify achievement of that objective. This is the end-task.
 - Identify tasks whose completion is required immediately preceding the end task.
 - Identify tasks immediately preceding each of those tasks.
 - Keep working backward in this fashion until you get to the starting tasks.
 - From the starting tasks, work forward one-by-one to validate succeeding tasks. This will identify any tasks that were overlooked in the backward pass.
5. Convert project plans into *buffered project plans* (stagger tasks to avoid resource conflicts within projects and insert integration and project buffers in the required places).

6. Challenge and refine assumptions (data) whenever the calculated project cycle time does not match the expected/desired result (see first item).
7. Share the final project plans¹⁵ with all task managers so that they understand their tasks (outputs, precedents, handoffs, etc.) as well as the overall plan.

Additional Tips

- A project plan is not a time reporting mechanism. The purpose of a project plan is to provide execution priorities and early warning signals.
- Noise factors like “lead and lag” dependencies or “fractional resources” should not be modeled.
- A project plan is not a to-do list. Tasks represent intermediate deliverables. If task managers or resources need a to-do list, these activities can be captured under the task as a checklist.
- A task is a chunk of work; definitive hand-offs of work characterize task scope. A task should not be broken down into several pieces just because it requires different resources at different times. However, tasks should be broken down to reflect handoffs among main resource types; that is, those resources that are required for most of the task time.
- Buffering policy should require projects to have a prescribed minimum amount of buffer before they can be accepted for execution. This will safeguard the buffering rule from project managers who might game the system by having smaller buffers, and from managers who might think buffers are unnecessary. Experience in small as well as large projects, and in one-off as well as repetitive projects, shows that about one-

third¹⁶ of a project's total buffered lead time should be buffer; shorter buffers make the priorities sensitive to even minor perturbations and longer buffers tend to delay managerial interventions.

Step 4: Establish Task Management

Task management is about assuring tasks are executed in the proper order of priority and with minimal interruptions, and monitoring remaining duration. Implementing and reinforcing this process is the key to sustained improvements in project performance.

Reporting Remaining Duration

Daily during execution, task managers estimate how much longer it will take to finish each of their tasks-in-progress. With this simple information, the amount of buffer consumed for the corresponding legs can be calculated and compared to the work completed in that leg. This information is then used to calculate task priorities and provide task managers with a report of all the current and upcoming tasks in order of priority, along with the rate of buffer consumption on the corresponding leg.

Tendency to procrastinate¹⁷ or not report early finishes is automatically curbed in this process. According to the head of engineering at a North American company, when "red" tasks were visible to all the concerned managers, task managers did not need much prodding to make or to report progress; every morning they would come in, follow up on their tasks to make sure progress was being made, and report the remaining duration.

Assigning Resources

Task managers assign resources to current tasks in order of priority. If resources are not enough to handle even the red tasks (tasks that have crossed the threshold of acceptable buffer consumption), overtime and other such decisions are implemented.

Preparing Tasks

After taking care of current tasks on their plate, task managers turn their attention to upcoming tasks. They ensure that all necessary preparations, such as getting approvals, drawings, materials, etc., are made so that tasks can be done without interruption as soon as the work of the preceding task is complete and available.

Organizations find it useful to formalize the responsibilities of front-line managers around the aforementioned aspects of task management.

Reminder: Do not pressure resources to meet planning estimates! Otherwise, you will soon be back to Square One.

Step 5: Implement Surrounding Processes

After reducing WIP and establishing task management, the benefits of synchronization will already be evident. Projects will be completing faster, firefighting and multitasking will be substantially less, and managers will feel more in control. However, the following surrounding processes are also needed to complete the picture.

Project Control

This is typically a formal weekly process to respond to uncertainties such as scope changes, technical problems, etc., that cannot be combated through routine task

management. If the rate of buffer consumption (percentage buffer consumed versus percentage work completed in the longest leg) is too high, then project managers know which legs of the project are in the “red.” They can then develop and execute recovery plans for those legs. Recovery plans can consist of run-of-the-mill items like scope adjustments and overtime as well as unique, even brilliant, solutions for specific situations.

Pipeline Control

While project managers can keep the buffers within their individual projects in control, it works only when just some projects are “red.” If most projects are running behind schedule, there is probably a more systemic or global issue at play that is affecting all projects in the pipeline. This is where senior managers step in and make global decisions like putting some projects temporarily on hold, reprioritizing projects, or authorizing across-the-board overtime.

For example, about a year after the initial implementation at an aircraft maintenance and repair depot, the number of red projects jumped from 30 to 70 percent. The underlying reason was a sudden increase in the sheet-metal work required on incoming aircraft. During the three months it took to ramp up sheet-metal capacity, the number of active aircraft in the sheet-metal department was reduced from four to three and the maximum allowed overtime was authorized. As a result, the duration of the sheet-metal phase came down from 65 days to 47 days; projects trended back to “green” and the required Throughput rate was achieved.

Pipeline Planning

During WIP reduction (Step 2), Execution is monitored to verify that the initially selected drum is still valid. If the original drum is getting starved for projects, then the real drum could be an earlier phase or an upstream resource; similarly, if queues build up downstream of the original drum, then the real drum could be a later phase or a downstream resource. The drum can be changed if such conditions persist.

In the event the drum changes, managers formally meet to reset project priorities and possibly revise due-date commitments. Then decisions about project priorities can be made routinely as new projects are undertaken or as business conditions change.

While actual decisions are made by managers in charge of the project operations (in consultation with other affected functions such as manufacturing, sales, and marketing), a dedicated “Master Scheduler” or “Pipeline Analyst” is typically required to provide analytical support.

Capacity Management

The loop is closed with a capacity management process that identifies and mitigates resource shortages. The required information comes from an aggregate database of project plans, which shows the total resource “load to capacity,” as well as buffer analysis, which identifies the resources that drive high-buffer consumption.

An important point is that the capacity of resources that are recovering buffers should be maintained or even increased (at least temporarily), even if it shows up as excess capacity in the “load to capacity” view.

In IT and engineering projects, for example, subject matter specialists do not have many explicit tasks in the project plans. A “load-to-capacity” view will show them as being 20 to 30 percent utilized. However, these specialists

are vital for recovering buffers. Keeping their planned workload at 20 to 30 percent is a good practice that ensures both project delivery and pipeline Throughput.

Step 6: Identify Opportunities for Continuous Improvement (POOGI)

Having put the Three Rules of Execution Management into practice, the holy grail of projects—how to prioritize improvement efforts—can now be pursued. Since almost every process in projects can be improved, it is essential to pinpoint and focus on those improvements that will have the biggest impact on global performance.

As is known, the most harm to lead times and Throughput is done by practices and resources that have the most impact on project buffers. Therefore, the logical way to prioritize improvement efforts is:

- Record reasons for delay in task completions.
- During buffer consumption calculations, identify the tasks that are affecting project buffers the most, and classify the corresponding reasons for delays.
- Do a Pareto analysis of the reasons for delays across all projects and address the top reasons.

Organizations that have focused and prioritized their improvement efforts in this manner have achieved even shorter cycle times and much higher Throughput than they achieved during the initial implementation.

Step 7: (When Applicable) Use Superior Delivery as a Competitive Advantage to Win More Business

When most competitors don't deliver their projects on time, and late delivery has a big effect on their clients, reliable due-date performance can give companies a competitive edge. Some companies in engineer-to-order manufacturing, after stabilizing Execution, have been able to win more clients by coupling their offers with large penalties for delays.

While completing projects early is not always relevant for the client, in some cases it is critical. For example, the U.S. Air Force captive MRO facilities improved their service and value by offering faster turnaround on aircraft that were in high demand.

Similarly, a supplier of equipment for power plants was on the critical path of projects to set up those plants. This supplier was able to increase its win rates without offering price concessions by promising and delivering shorter lead times.

Lessons Learned

Following are some of the key lessons¹⁸ drawn and shared by hundreds of managers who have implemented Critical Chain.

Performance Gains Come from Managing Differently, Not Better Planning and Visibility

While good plans are essential and the rate of buffer consumption status is an effective way to monitor projects, increasing the rate of execution requires changing the way execution is managed.

Projects can be planned better even without Critical Chain, and the rate of buffer consumption provides similar information about project status as a comparison of actual time-lines against the baseline. If only better planning and

visibility are what is required, they can very well be achieved with traditional methods.

However, unless WIP is reduced, task-level measurements are abandoned, projects are planned with shorter cycle times, and buffer-based priorities are followed, execution priorities will not be synchronized and projects will not be done faster. Nor will Throughput be increased.

Implement All of the Three Rules

Experience over the years has shown that the Three Rules of Critical Chain must all go together (see [Fig. 4-2](#)). Not implementing any one of them only shows up as lack of results or resistance to change. For example, organizations doing multiple projects with shared resources might be tempted to implement Critical Chain one project at a time. They ignore the pipelining rule. In a shared resource environment when WIP is not lowered, conflicts for resources continue. Priorities cannot be followed, buffers are consumed, and commitments are missed. Very quickly, faith in Critical Chain is lost.

Many times organizations aim just to gain control without increasing speed and Throughput. They compromise the buffering rule (for example, cycle times are not cut, but buffers are added). When cycle times are not cut, pipelining is compromised because long cycle times mean high WIP. When WIP is not lowered, Buffer Management cannot be done. The entire system falls apart.

Some managers compromise Buffer Management because they feel this is micromanagement. However, without people working to a single priority system and without timely interventions, buffers are wasted. This creates a sense that shorter cycle times were unrealistic. Eventually the organization reverts to its old ways (high WIP, safeties

embedded inside individual tasks, and ad hoc priorities in Execution).

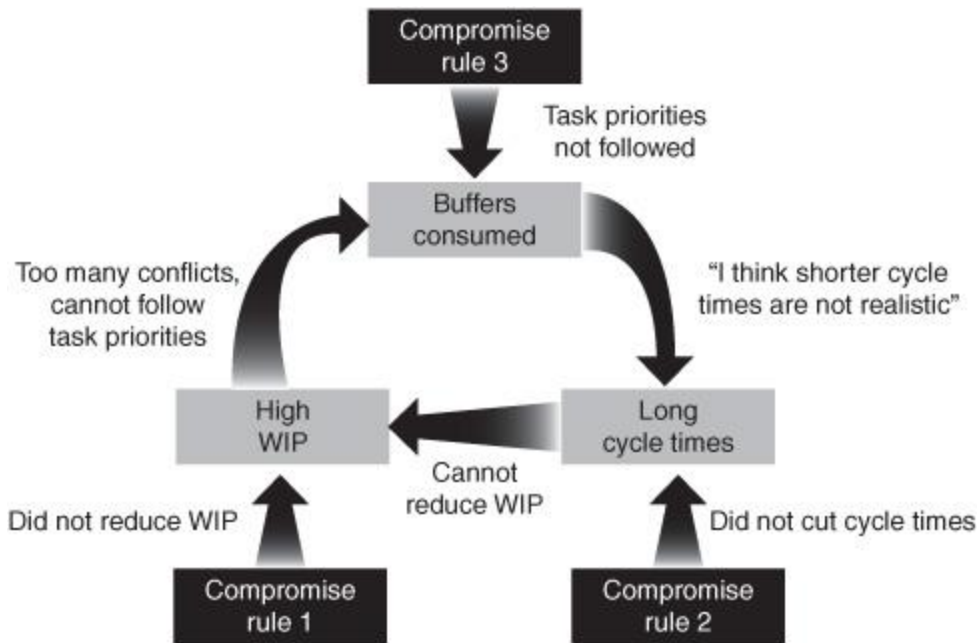


FIGURE 4-2 Why implement all the Three Rules.

Top Managers Must Play an Active Role

Mere sponsorship by top managers is not enough. Even though the top managers' role is typically to set policies and make planning-time decisions (project execution is delegated to middle and front-line managers), in successful implementations the top managers take on a more active role for the first 6 to 12 months.

The first reason is that middle managers and front-line managers encounter policy obstacles that they do not even know can be removed. Only senior managers can identify and eliminate those policy obstacles. For example, middle managers frequently assume that project starts cannot be staggered because clients will not buy-in; however, when the matter is brought up to top management, they are often

willing to explain personally to their clients the benefits of pipelining projects. The CEO of one medium-sized manufacturer of industrial equipment even undertook a tour of customers around the world to explain pipelining and get their buy-in.

Second, managing buffers takes time to become a habit. It is only human to revert to old ways as soon as there is a minor hiccup. Close oversight by top management is necessary until managing buffers becomes second nature (“constantly peering over the shoulders” as an engineering manager from one company put it). The leadership in a U.S. Air Force Logistics Center went on daily rounds and for three months personally got involved in resolving issues.

Finally, outsiders can teach concepts. However, how to manage differently is better “taught” by top managers. For example, officers of senior rank in military organizations and “C” level executives of multibillion dollar companies have personally taught and coached their middle and front-line managers in the principles and practices of Buffer Management.

Actively Manage the Buffers

Buffer reports provide an accurate status of Execution. However, merely communicating status is not where the advantage of buffer reports is. The power of Buffer Management comes into play only when used by managers to respond actively to uncertainties. Here is how buffers are managed at various levels in an organization:

- **Task managers**—In contrast with traditional project management, the advantage of Critical Chain in execution is at task level because that is where the work is done. All organizations implementing Critical Chain, ranging from tens to thousands of people working on

projects (whether they do research, engineering, or manufacturing projects), have realized the importance of task management. Talking about its implementation at a fashion garments supplier in Australia, the responsible person observed: “It is quite simple. You update your tasks, follow priorities, and get the work done.” According to the engineering director of a home appliances company, “Setting processes and guidelines for Task Management is the key.” Another successful adopter from a submarine maintenance facility put it as, “The supervisors look at their task list and allocate resources based on priority. It is that straightforward.”

- **Project managers**—According to a provider of telecommunication switches, there was tendency in their implementation to use project review meetings only to explain “red” buffers. Only when the division managers started expecting actions for *recovering* buffers did the projects begin being brought on track.
- **Resource managers**—At a provider of IT applications, resource managers initially did not see a role for themselves in managing execution. However, after Buffer Management was in place, it became evident to them how to anticipate and prevent resource bottlenecks rather than scrambling for resources *post facto* (one of the outputs of Buffer Management calculations is an accurate list of upcoming tasks in every department and corresponding workload). Moreover, their earlier resistance evaporated.

Frequently Asked Questions

Following are some additional implementation-related questions and answers drawn from field experience.

Can Critical Chain be implemented without basic project management in place first?

It is worthwhile debunking the myth that Critical Chain might be too advanced; that project management basics have to be well in place before Critical Chain can be implemented. It has been observed that many of the so-called “basics” actually propagated and reinforced the old ways of running projects. Organizations that were mature in “basics” actually had to let go of some of the practices they had acquired; for example, making detailed project plans and issuing precise task schedules. Organizations that did not have the required fundamentals such as good project plans or management structure could quickly establish them as part of their implementation. The effort is not on establishing the “basics,” but on implementing Critical Chain itself.

Should a pilot be run before a full rollout of Critical Chain?

Not necessarily. The application of Critical Chain is now well understood for a wide range of project types. A pilot is not needed if experienced implementers who have successfully performed similar implementations elsewhere are engaged.

If implementing alone and for the first time, without help from experienced implementers, a pilot might help¹⁹ in understanding the implications of all Three Rules.

If external help is available but does not have experience in a relevant business or operational environment, pilots may be advisable for the same reason. However, it is important to set clear objectives for the pilot (e.g., what specific changes to test and what effects to measure) and structure a pilot accordingly.

What about cultural and behavioral changes?

Organizational culture and peoples' behaviors cannot change *before* results happen. The culture and behaviors under Critical Chain are undeniably quite different from traditional culture and behaviors. At the same time, culture and behaviors are broad and nebulous terms; if not careful, they can become a smoke screen to hide real implementation issues.

More importantly, culture and behaviors stem from how you manage. Change the rules and associated policies and measurements, and the culture and behaviors will begin to change as well. Results that come from new rules will only accelerate those changes.

What is the role of software in Critical Chain?

The main role of Critical Chain software is enabling and leveraging Buffer Management.

Many project planning tools can create satisfactorily buffered project plans (albeit with a lot of manual effort), and even spreadsheets can adequately plan pipelines. However, for Buffer Management, specialized software components are required:

- A computational engine to monitor buffers and calculate priorities
- A central database to collect the inputs and outputs of buffer management
- A Web-based platform to capture and disseminate information in real time across the enterprise

Specialized software can also play a significant role in sustainment by monitoring and reporting the results as well as adherence to the Three Rules.

Is a Project Management Office (PMO) needed with Critical Chain?

While a specialized group is needed to support a management system based on Critical Chain, it is quite different in nature from a traditional PMO. Whereas the traditional PMO is mostly about planning and reporting, often with an explicit focus on improving and enforcing task estimates, a Critical Chain support group is about facilitating synchronized execution. Its role is applying and enforcing the Three Rules by helping:

- senior staff maintain low WIP, create pipeline and capacity based on business goals, slot new projects into the pipeline, and monitor results and adherence to the Critical Chain Rules;
- project managers create properly buffered project plans;
- task managers follow priorities;
- identify and capitalize on opportunities for continuous improvement; and
- train and coach new managers.

To communicate and reinforce a clear change in focus, it is probably appropriate to term the Critical Chain support group as *Execution Management Office* (EMO). This group should be professionally knowledgeable in Critical Chain Rules and practices, and expert in execution management software.

How is non-project work handled with Critical Chain?

Up to 10 to 50 percent of work that a typical project-based operation has does not come from projects. Examples of such work include sales support, field support, and special

tasks that cannot be classified as projects. All such work potentially interferes with following buffer-based priorities for project work. To make matters worse, non-project work often does not go through a central coordination and control point, or gate; it just lands on people's desks.

When non-project work is little (~10 to 15 percent of the total workload for a set of resources), a practical solution is to establish a central gating and dispatching mechanism. Emergency work is immediately assigned, preferably to those people who are not working on red tasks, while other work is assigned to people as they finish their project tasks.

If non-project work is substantial (more than 20 percent of the total workload for a set of resources), it is best to dedicate capacity for it. Otherwise, not only will it be difficult to follow buffer-based priorities for project work, but non-project work will suffer as well. If it is important to give everyone a chance to perform project as well as non-project work, a rotating pool can be established whereby people are assigned to non-project work for only a few weeks at a stretch.

Should the scope of a Critical Chain implementation include vendors and subcontractors?

If vendors supply long lead-time items, and procurement of those items is on the projects' critical path, the improvement in project cycle times may be limited if vendors are not included in the implementation. Organizations can still achieve the full potential increase in Throughput of internal resources (typically 20 to 25 percent), but only 10 to 15 percent reduction in overall cycle times. Achieving greater reduction in cycle times requires offering vendors an incentive for faster supply, and perhaps implementing Critical Chain (or Drum-Buffer-Rope) in the vendors' operations.

If subcontractors perform a significant amount of work for the project, the improvement gains in Throughput as well as cycle time may be limited if they are not involved. If proper incentives are provided, subcontractors can be persuaded to execute their work in accordance with the Three Rules of Critical Chain—to the benefit of both parties.

How does Critical Chain improve quality?

Critical Chain helps improve quality by cutting down firefighting and multitasking and by creating time at the beginning of a project for full kitting. Moreover, metered release of projects checks the temptation to start them before fully defining the requirements, which minimizes later changes and the rework, errors, and multitasking that emanate from them.

Critical Chain seems to be all about timelines; what about controlling costs?

Of course, costs of a project cannot be managed without regard to project benefits. It is sometimes possible that the benefits far outweigh the costs of doing projects faster. In most other cases, costs can be a relevant concern.

There are two viewpoints about timelines and costs. One viewpoint is that they are in conflict—shortening timelines costs more money. The other is that the longer projects take, the more they cost, so there is no need to worry about costs as long as projects finish faster. However, both these assertions are only partially correct.

Many adopters of Critical Chain have found that project costs can be divided into three categories:

1. **Costs of Capacity:** Costs of people, equipment, and facilities fall into this category. The faster projects are

done, the earlier the capacity is freed up, and the capacity costs incurred by individual projects are lower. This applies if projects are done faster without increasing the rate of expenditure on resources (e.g., by expediting, spending overtime, etc.). If resources are fixed, and these resources can complete more projects within a fixed time, then the average cost of the completed projects declines. Similarly, if projects are delayed, the (assigned) cost of individual projects could increase.

2. **Costs of Purchased Items:** Costs of material and components, and firm-fixed price work done by subcontractors, fall into this category. Such costs likely will not change with project duration, except if supplies are expedited. Such costs are best controlled with traditional methods, within the framework of Critical Chain policies and practices.
3. **Costs of Expediting:** The exceptions to the previous characterizations are the costs that can be incurred to recover buffers; this includes costs of extra capacity as well as paying premium prices to expedite materials, buying materials that are more expensive, or transporting materials using faster modes of transportation, and the like. Of course, such costs should be incurred only if the benefits of improved delivery or reduced risk outweigh the costs.

Potential conflict between timelines and expediting costs can be mitigated by recognizing that buffer recovery actions at additional expense may be required during execution. A useful and prudent practice is to set aside monies to help recover buffers as necessary. This “budget reserve” is part of the total budget, not in addition to it. Experience says that 10 to 20 percent of the total budget is appropriate for

“budget reserve,” and setting it aside upfront helps prevent cost overruns while delivering projects on time.

Do we need project-level budgets in multi-project operations?

Since costs of capacity in multi-project operations are not incurred project-by-project but in aggregate, it is not necessary to budget these costs at the project level. An aggregate budget is generally sufficient for controlling the costs of capacity. However, a project-level budget may be helpful for managing the costs of purchased items. In addition, organizations might still need project-level budgets for reporting to their customers and for financial accounting purposes.

Does Critical Chain work with Earned Value Reporting?

It is quite straightforward. Organizations contractually obligated to report Earned Value metrics continue doing so even after they implement Critical Chain. However, they do not use CPI²⁰ or SPI²¹ to measure Execution and drive execution priorities. They use Buffer Management for that.

How does Critical Chain work with Lean?

Lean has three well-known elements: *Kanban*, which is about synchronizing execution priorities and tying them to the actual demand; *Flow Lines*, which are an alternative to *Kanban*; and *Kaizen*, which is about a process of continuous improvement.

Kanban normally does not apply to projects. *Flow Lines* have been tried in project-based manufacturing but without much success. The reason is that Flow Lines require reliable estimates of time and effort required to do a task, which are

not possible in projects. In short, there is no alternative to Critical Chain for synchronizing project execution.

The difficulty with *Kaizen* in projects is that on the one hand, almost everything can be improved, and on the other, most local improvements do not translate into better project performance. Buffer diagnostics can enable *Kaizen* by helping to isolate and prioritize meaningful improvement opportunities.

In other words, Critical Chain *is* Lean for projects.

What are the likely causes of failure in implementing Critical Chain?

The implementation process presented in this chapter is the fruit of over 200 enterprise-level implementations since 1999. Before this process was developed, roughly only one-third of the adopters realized significant improvements in project speed and Throughput; another one-third experienced marginal improvements (projects in control and on time); while one-third of the implementations failed to take off. Since introducing this process, the success rates have been near-perfect. Significant improvements are realized every time this process has been followed. However, the following points of failure can occur, and prevent an adopter from following the prescribed steps and enjoying consequent benefits:

- Undertaking an implementation without a business imperative.
- Top management not accepting or setting sufficiently ambitious improvement goals, or delegating the implementation to a staff function like a PMO. Critical Chain inherently involves changing the rules of managing Execution and performing at a higher level, not about planning and tracking projects differently.

- Not changing the policies and measurements that conflict with the Three Rules; local (task-level) schedules and measurements are the biggest culprit.
- Inability of the implementation team to apply the Three Rules to the environment under consideration. The most difficult parts are applying the pipelining rule, building good project plans (see Step 4), and designing and establishing task management.
- Activating Buffer Management reports but not following through with coaching and mentoring of front-line managers in actively managing the buffers.

If the business case for Critical Chain is strong, and if any of the other failures mentioned occur, the reason is either lack of implementation skills or inadequate leadership.

Summary

Critical Chain works because it solves the real problem caused by uncertainties that are inherent to projects. It recognizes that while uncertainties can be somewhat lessened through better planning, they cannot be significantly reduced or eliminated. Therefore, Critical Chain curbs the immediate and most devastating effect of project uncertainties—unsynchronized priorities. The Three Rules provide an assured basis for coordinating projects' tasks and resources to achieve optimal performance.

Second, getting results from Critical Chain pragmatically focus on translating these explicit Rules into practical procedures *before* trying to change behaviors and culture. Experience has consistently shown that practical procedures and robust buy-in of managers to the Rules are enough to get results quickly. Management buy-in is solidified by quickly achieving specific improvement targets based on

real business needs. When Critical Chain Rules are also then embedded into management policies, management processes, and management information systems, organizations get as close to long-lasting and self-perpetuating results, culture, and behaviors as is possible in “human systems.”

Finally, there is no alternative to strong leadership—either for getting initial results or for ongoing improvements. Only top managers can change the old rules and preserve the new Rules for managing Execution. Only top managers can set appropriately ambitious goals for the organization. Any other assumption is folly and leads to failure.

References

Goldratt, E. M. 1997. *Critical Chain*. Great Barrington, MA: North River Press.

Goldratt, E. M. 2008. *The Goldratt Webcast Program on Project Management: Sessions 1-5*. (Video series: 5 sessions) United Kingdom: Goldratt Marketing Group.

Goldratt, E. M. and Goldratt, A. (R). 2003. *TOC Insights: Insights into project management and engineering*. Bedford, UK: Goldratt Marketing Group.

Realization. 2010. Case Studies. Accessed March 30, 2010 at: http://www.realization.com/case_studies.html

Realization. 2010. Critical Chain Results. Accessed March 30, 2010 at: <http://www.realization.com/customers.html>

Realization. 2010. Lessons Learned. Accessed March 30, 2010 at:

http://www.realization.com/projectflow/lessons_learned.html

Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>

About the Author

Realization Technologies, Inc. is a leading provider of Project Execution Management solutions based on Critical Chain. It serves clients in a wide range of industries around the world, including organizations like ABB, Alcatel-Lucent, Amdocs, Boeing, CNAT (Centrales Nucleares Almaraz-Trillo), Delta Air Lines, Dr. Reddy's Laboratories, Hamilton Beach Brands, Iberdola, Larsen & Toubro, Medtronic, Procter & Gamble, TATA Steel, Vale, Votorantim, the U.S. Air Force, the U.S. Army, and the U.S. Navy.

CHAPTER 5

Making Change Stick¹

Rob Newbold

One day, the Master caught his favorite apprentice stealing. He said angrily, "I will not train a thief. Go, and return when you have changed."

The apprentice, feeling very ashamed, spent the day walking around the village, thinking about his life and his behavior. He returned that evening, saying, "Master, I have spent the day reflecting on who I am and what I would like to be. I believe I will act differently in the future. With all my heart I wish to return and continue to be your apprentice."

The Master replied, "Realizing that you need to change is not change. Go, and return when you have changed."

The apprentice, much dismayed, set out again. This time he traveled to a nearby city, working from time to time to support himself. After two weeks he returned, saying, "Master, I have spent two weeks working and learning and have never once been tempted to steal. I know I will act differently in the future. With all my heart I wish to return and continue to be your apprentice."

The Master replied, "Trying new things is not change. Go, and return when you have changed."

So the apprentice set out for a third time, traveling the country far and wide, learning skills and seeing wonders of which he had never dreamed. After a year he found himself near his original village and stopped in to visit his old Master. He said, "Master, I have traveled the world and seen many wonderful things. I am happy to see you, but being your apprentice is no longer my heart's desire."

The Master smiled and said, "You are welcome to stay as long as you wish."

Introduction

Real change, the kind required to significantly improve organizational performance, is not about realizing we need to change. It is not about trying a few things. It is about changing our habits, the habits we use without thinking as we respond to daily situations. When we implement Critical Chain scheduling, we want people to do certain things without having to think about them. For example:

Copyright © 2010 by Rob Newbold.

- Perform work as a *relay race*² ("*get it, work it, move it*"), not a train schedule.
- Assess actions through their impact on the global project or portfolio picture, not through their impact on task due dates or individual productivity.
- Treat commitments as ranges of time, not points in time.

Use of these concepts represents a real shift of paradigms for most organizations. Until new habits are part of the organization's DNA and the old habits are gone, people have to weigh alternatives and consider multiple approaches. They have to think. Meanwhile, the old approach continues to be an easy option, so backsliding is common. Until the DNA has changed and new habits formed, the change process is not complete.

In this chapter, I explain the approach to organizational change developed and refined at ProChain Solutions over the last 12 years as we helped organizations of all sizes implement Critical Chain Project Management. First, I will analyze the nature of the problem and the root causes behind change not sticking. Then I will discuss a solution, the Cycle of Results (CORE), and how this solution can be used to address the root causes. Finally, I will describe how CORE can be applied to the implementation of Critical Chain scheduling.

The Uptake Problem

A major reason that organizations are unwilling to take on major change initiatives is what I call the "Uptake Problem." Implementations have trouble getting off the ground; when they do, they don't produce to the level people believe is possible; and even when significant benefits are produced, backsliding can, over time, put an implementation in jeopardy. Many times the Uptake Problem is explained simply by saying, "Change is difficult" or "We're not good at change."³

The Uptake Problem is readily acknowledged across many types of implementations, but is very difficult to quantify. Experts and companies seldom have an incentive to reveal negative data, so we see only one side of the picture. When we find and accept negative data, the extent of the Uptake

Problem—difficulties getting going, the extent to which improvements continue, etc.—is difficult or impossible to analyze. Even the definition of success will tend to vary by time and organization. There are tidbits that reference the problem, but never a full meal.

- Yearly surveys from the Lean Enterprise Institute indicate that backsliding is a perennial problem. (Lean Enterprise Institute, 2008, 1)
- “Although individual lean concepts and tools are easy to understand, to be truly successful in the application of these concepts and tools, the majority of the organization must change the way it looks at work. . . . And so far, the vast majority of the organizations that start on the lean transformation journey are not successful at making this transition.” (Koenigsaecker, 2009, 79)
- “Statistics from 150+ implementations . . . 15% of the implementations failed to take hold[,] despite initial successes[;] 15% of the implementations failed to even take off.” (Gupta, 2005, 3)
- “In practice (in our experience) most [Critical Chain] implementations have failed after the person driving the process has moved on.” (Retief, 2009, 1)
- Hobbs and Aubry found that 42 percent of Program Management Offices (PMOs) have had their relevance or even existence seriously questioned in recent years, leading them to believe that, “... about half of organizations are critical enough of PMOs to decide not to implement one or to seriously consider shutting theirs down if they already have one.” (2006, 13).
- Frequent anecdotal evidence suggests that the Uptake Problem is significant with any major change initiative, including *Theory of Constraints* (TOC), Enterprise

Resource Planning, Enterprise Project Management, Lean, and Six Sigma.

ProChain's experience, gained over the course of 12 years observing our clients and the clients of others implement Critical Chain, confirms that the Uptake Problem is real and pervasive. We have found that:

- The Uptake Problem is more severe with larger projects, larger organizations, and organizations that perform projects involving significant uncertainty (e.g., research and development).
- The immediate value of Critical Chain to project managers is such that individual project managers, once trained, will often attempt to continue to use it whether or not the organization embraces it.
- People (and organizations) who take the perspective that Critical Chain is a toolset, rather than a significant change process, are unlikely to maintain successes over the long-term.
- There is a direct correlation between implementation success and willingness to adopt the CORE concepts described in this chapter. For example, every company that has started a ProChain rollout within the last five years has continued increasing their use of and value obtained from Critical Chain over time.⁴

Before we can fix the Uptake Problem, we need to understand it. The following analysis follows the *Current Reality Tree* (CRT) shown in [Fig. 5-1](#) through [Fig. 5-3](#). A CRT is a *tool to pinpoint common causes responsible for many effects*.⁵ Read the boxes in the tree in numeric sequence. Boxes with no arrows leading into them are “root causes”;

they should be examined for validity. Other boxes should be read following the arrows, using if-then logic. When multiple arrows go through an ellipse, read “and.” For example, starting at the bottom of [Fig. 5-1](#): “*if* (1) sometimes people lack urgency to change to a promising new technology *and* (2) there is some level of interest in the new technology, *then* (3) there are half-hearted attempts to employ the new technology.” Boxes that have already appeared in an earlier figure are shown in a light shade of gray.

In order to make this discussion as concrete as possible, imagine you are employed by a large company, Widgets, Inc. (WI), as a project manager for new product development. WI designs and manufactures (of course) widgets—big ones, little ones, all kinds. I have included some narrative to describe the CRT logic as it applies to WI. I have added the associated box numbers from the CRT into the narrative in parentheses.

No Urgency to Change

Suppose, to start, that WI, as a whole, isn’t experiencing significant pain with its projects, meaning there is little urgency to change even to a promising technology (1). New products are coming out of the hopper, the system doesn’t look broken, so there is no urgency to fix it. Despite that, you as a project manager are interested in implementing Critical Chain scheduling because you recognize that it will have value for you (2). How are you likely to fare?

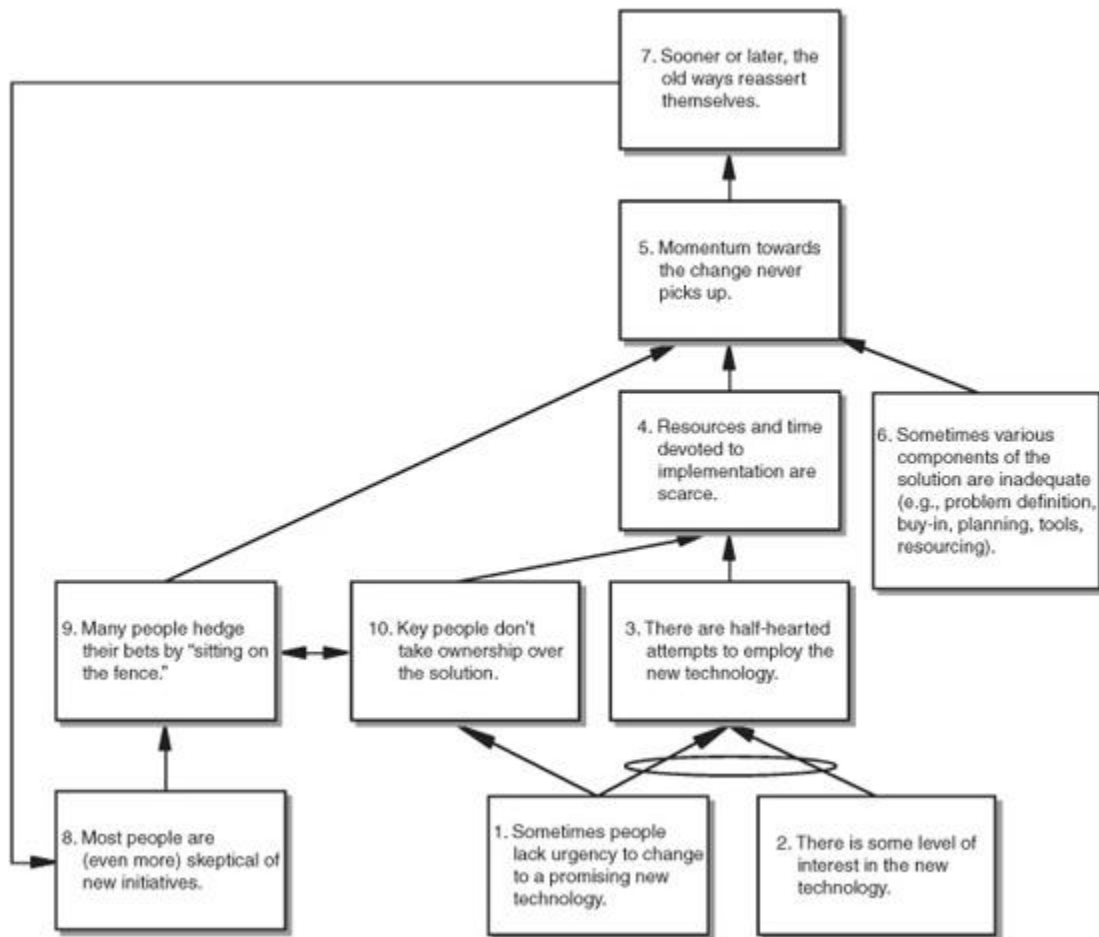


FIGURE 5-1 No urgency to change.

You may gather support from some like-minded individuals, but Critical Chain schedules will only meet with half-hearted interest (3); people have too many things to do that are more important. Needed resources and time will be scarce (4). Consequently, while you may use Critical Chain for your projects and people may or may not express interest, the momentum never builds (5). Of course, if basic components of your solution are deficient (6), your chances of building long-term momentum will be even worse.

Since Critical Chain (like other TOC applications) requires the synchronization of many people to be fully effective over the long-term, and since the momentum is not building, your implementation cannot take off. Not enough people are

synchronized; the old DNA is not being replaced. Eventually, as enthusiasm wears off or people move on to other positions or companies, the old ways reassert themselves (7).⁶

There is also a loop that makes things worse. People have often seen initiatives fail; these failures tend to make people skeptical of new initiatives (8). Why bother rearranging the deck chairs on the Titanic? These bad experiences, and stories of bad experiences, often cause people to take a wait-and-see attitude toward change (9). This attitude by itself reduces the momentum toward change (5).

There is one other problem that frequently exists and is made worse by a lack of urgency (1): key people—typically mid- and high-level managers—will not take ownership over the solution (10). Without their ownership, resources and time remain scarce (4). Note the twoway link between boxes 9 and 10. When key people do not take ownership, others will assume that it is all right to sit on the fence. The more people there are sitting on the fence, the more key people are likely to avoid taking ownership.

Throughout all this, people working inside WI will have real trouble understanding what happened. They may say things like:

- Our culture wasn't ready.
- Our focus changed.
- We never got the management support we needed.
- We just couldn't execute.

Their thinking is governed by the general skepticism of new initiatives (8).

The Silver Bullet

Moving on to [Fig. 5-2](#), suppose that, due to an aggressive new competitor, senior management at WI starts to believe there is an urgent need to reduce cycle times without increasing costs (11). Senior leaders therefore put their weight behind a Critical Chain initiative to reduce cycle times (12). They also make sure to have in place all the basic components of a good solution (13).

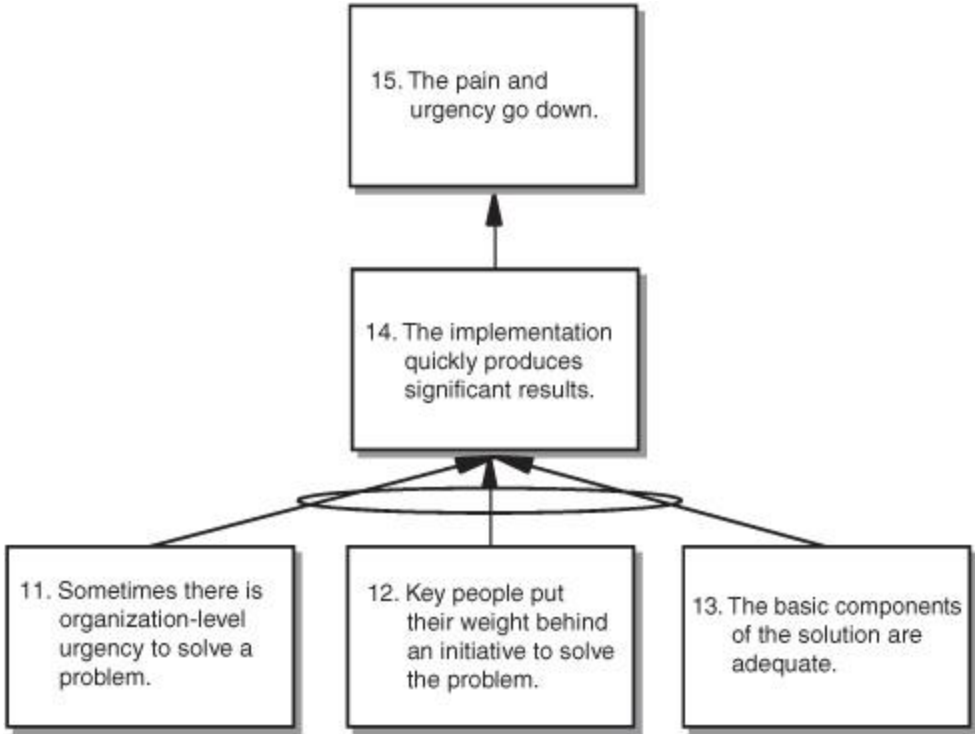


FIGURE 5-2 The silver bullet.

Stage	Task
Preparing	Study and analyze the organization.
	Work with key players to plan the implementation.
	Install and configure software and other tools.
Teaching	Give practical training, potentially to many different types of individuals, for task, project, and portfolio planning and execution.
	Provide mentoring, especially for project and functional managers.
Sustaining	Initiate formal processes for certifying the quality of internal experts
	Create a PMO to manage methodology and quality.
	Adapt Human Resources processes to reflect the importance of project management.

TABLE 5-1 Generic Critical Chain Implementation Steps

[Table 5-1](#) shows generic, high-level components of an implementation plan WI might use.

Each of these components is important and worthy of its own discussion. We have seen implementations in which a lack of any one of them has proven fatal. However, assuming WI does all these things, they are likely to achieve rapid and significant successes with their implementation (14). They have relieved the pain and urgency (15). In short, WI's Critical Chain solution has proven to be a real silver bullet.

Negative Branches

A negative branch (Scheinopf, 2000, 117) is *something bad that happens because of trying to do something good*. It is the road to hell, paved with good intentions. For example, there are likely to be negative branches associated with giving money to a drug addict. We have seen a few negative branches arise repeatedly in implementations, even when the implementations produced real benefits; they are shown in [Fig. 5-3](#).

Let's assume that [Fig. 5-2](#) is completely valid, and by virtue of WI's quick successes with Critical Chain their cycle time pain has gone way down. The consultants leave and chalk up a success and everyone is happy.

Unfortunately, it is not common practice to set clear, realistic expectations of the likely value of an initiative (16). It is also uncommon to communicate that value once it has been achieved (17). As a result, many key people do not fully appreciate the value provided by the initiative (18).

Two problems occur when key people do not appreciate the value. First, if they are looking at costs versus benefits, they will get a skewed picture, especially keeping in mind that different people will have different perceptions of value. The benefits may not appear adequate relative to the costs (19). Therefore, while WI's CEO may appreciate the tremendous benefits from their reduced cycle times, a functional manager may only see that his job is less important because his firefighting skills have become irrelevant. That functional manager will be less willing to lend his support to the initiative (4).

The second problem is that the success of the initiative means that people see the problem as "solved." Unfortunately, there are still associated costs: internal experts, consulting support, licensing fees, and so on. Who wants to continue to pour time and money into a problem that is solved (20)? Often, they won't (4). In addition, do not forget the fence-sitting loop from [Fig. 5-1](#) (box 8): Repeated failure makes people more and more skeptical that success is possible.

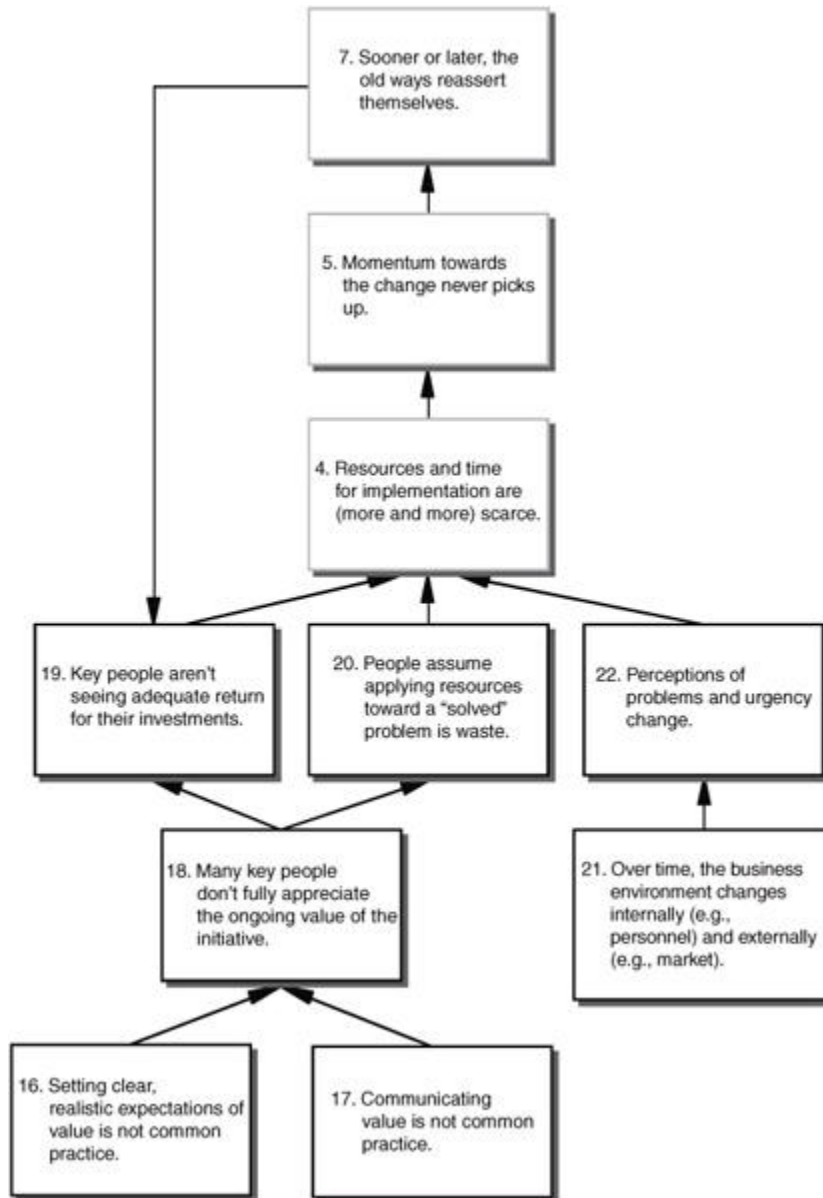


FIGURE 5-3 Negative branches.

There is one more ticking time bomb: business environments change over time (21). We have seen the replacement of a senior executive produce a complete change in the focus of an organization. We have also seen something as simple as a market downturn result in drastic cost cutting. In other words, the perception of what

problems exist, and their urgency, changes over time (22). Again, support disappears.

All these negative branches will eventually make WI's Critical Chain initiative a legitimate target for cost reductions. Support personnel have to do more with less; management may even decide eventually to eliminate the PMO.⁷

The inescapable conclusion is that even the most successful and apparently well-managed initiatives are under threat. It is no surprise that people are skeptical that change can really happen.

Root Causes

[Figures 5-1](#) to [5-3](#) suggest some root causes that drive the failure of change initiatives:

1. Lack of urgency (box 1)
2. An inadequate solution (box 6), including:
 - The problem, appropriate solution, or needed results are poorly defined.
 - Buy-in of key players is inadequate.
 - The implementation plan does not address major obstacles.
 - Insufficient resources are applied to the solution.
3. Lack of ownership in the solution (box 10)
4. Unwillingness to set clear expectations of value (box 16)
5. Inability to communicate value (box 17)

6. Changes in the business environment (box 21)

The Cycle of Results, presented in the next section, addresses the first five of these root causes. The sixth, changes to the business environment, looks like an inevitable result of doing business. However, it has a couple of important implications. First, in the midst of a change effort, management should be careful about the additional changes it can control. We have often seen new change initiatives taken on before old ones have been assimilated. We also often see key personnel moved around without regard to the impact on initiatives. Management should minimize these changes.

Second, the inevitability of changes in the business world implies that there is a finite window for any implementation to take hold. If you cannot set up the appropriate processes before the next earthquake, your initiative will eventually be in trouble.

The Cycle of Results (CORE)

The implementation feedback system used by ProChain to address the Uptake Problem is called the Cycle of Results™ (Newbold 2008, [Chapter 15](#); also referred to as CORE). It addresses the first five root causes of failure from the previous section to create a process that builds trust. I define trust as *willingness to depend on someone or something, in a specific context*. The people implementing the solution must be willing to depend on that solution, continuously, into the future. People must believe that the perceived rewards will continue to outweigh the perceived costs.

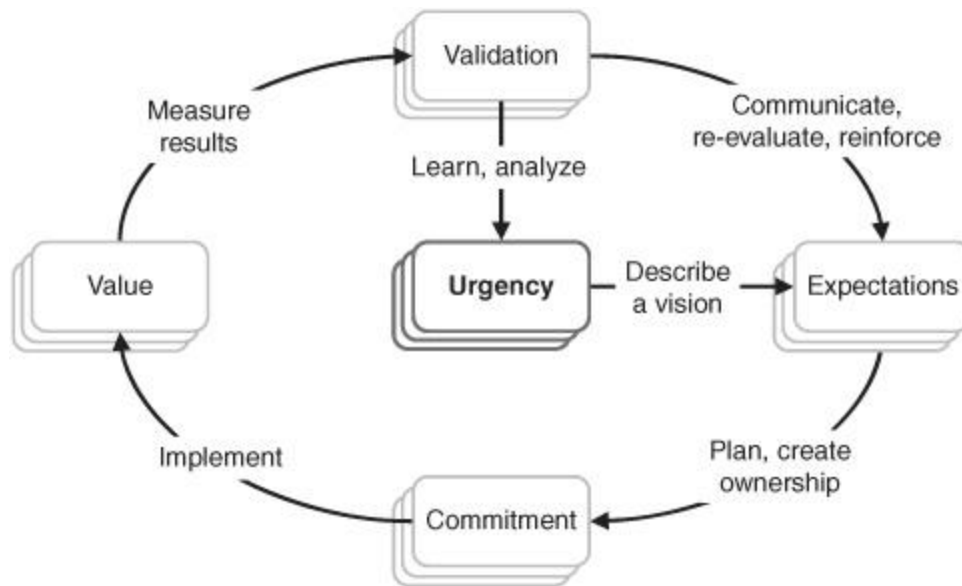


FIGURE 5-4 The Cycle of Results. (Copyright © 2008 by ProChain Solutions, Inc. Reprinted with permission.)

Basic Principles

[Figure 5-4](#) shows CORE pictorially. In this picture, the conditions or states achieved are in the boxes and the actions leading to those states are on the arrows. The boxes are stacked because the meaning of the states may be different for different people. Different people may feel different types or levels of urgency, have different perceptions of value, and so on.

The cycle starts with the actions leading into Urgency: *learn and analyze*. It is especially important to learn and analyze the urgency that people experience to change. Suppose you are a consultant and someone asks you to help him or her implement Critical Chain for a project. Do you immediately convene the project team, or do you first try to understand why the organization wants to implement it? As discussed later with implementation planning, we may need

to take many actions in order to understand what urgency people feel today and what urgency we need them to feel.

Because of its importance, *Urgency* is in the center. It is a necessary condition for meaningful change. Urgency may be different for each person, so for example a senior leader may experience urgency to improve revenues, a project manager may experience urgency to deliver a project more quickly, and a worker may experience urgency to finish a specific task. You will need to understand the urgency for different individuals, because they will not respond to urgency they do not feel. Very often, I have heard people say that they believed their Critical Chain implementation had urgency “because my boss says so.” If it is their boss’s urgency, it is not theirs. If it is not theirs, they do not really feel it. If I see on television that a building is on fire, I will feel badly for the people inside, but I probably won’t run out my door to escape the flames.

How do we combine these different feelings of urgency into a whole that is synchronized around the needed improvement initiative? We need to *describe a vision* for the implementation, a vision that connects what the company does (and why people want to work there) with the benefits they should expect from the implementation. For example, a simple vision for WI might be, “We will improve our customers’ lives and our ability to compete for their business by getting them the new widget technologies they need when they need it.”

This vision should be described to the different people in the organization in their terms, in order both to set *Expectations* for the implementation and to tie the expectations to people’s individual sense of urgency. For a Critical Chain implementation, senior leadership will need to understand the strategic and bottom-line implications. Project managers will need to understand that they will be able to focus on high-impact actions. Financial people will need to understand the financial ramifications on

predictability and resource allocation. Individual contributors will need to understand that they will be allowed to focus. And so on.

Sometimes it seems that senior people set up initiatives and make promises without providing the wherewithal to actually make things happen. To avoid that, we recommend a significant *planning effort*, beyond any generic plans that may already exist. This has two benefits. First, it makes sure that specifics of the environment are taken into account so that they do not cause problems later. For example, the organization's structure will likely have a significant impact on the sequence of implementation activities. Second, people are much more likely to take ownership over things that they have had influence on. That is true whether you are creating gardens, businesses, or implementation plans. When it comes to organizational change, people in the organization are either part of the problem or part of the solution.

The planning process, along with related activities such as interviews with stakeholders, helps to build an initial *Commitment* to move forward. It won't get everyone off the fence, but it will help start the key individuals moving.⁸ With that in mind, we will often start a major implementation with both a senior leadership group or Steering Team and a slightly lower-level Implementation Team. The Steering Team gives planning advice and approval, the Implementation Team does the more detailed planning. They all have a say in what happens.

After some level of commitment comes *implementation work*, in order to create *Value*. Creation of value for all the key stakeholders seems to be straightforward for Critical Chain implementations because so many kinds of value can be created. [Table 5-2](#) presents some examples of the benefits of a full enterprise implementation of Critical Chain to the different players.

We have seen these types of value repeatedly. However, this brings us to the top of the cycle: how do we know that the value was achieved? It must be measured. [Table 5-2](#) includes some sample implementation *measurements*.⁹ One that we have found to be of great value but not commonly used is the last in the table: checklists to determine process adherence. During a weekly buffer update meeting, one would expect certain topics to be discussed: buffer consumption, recovery plans, key tasks, and so on. During functional staff meetings, one would expect discussions about how to work one task at a time. Why not use a checklist to track whether these things are happening? We have found this to be a great way to learn where help is needed.

It is never enough just to measure. The measurements must be validated against different people's expectations. In addition, once that *Validation* has taken place, the results must be *communicated* with key stakeholders. If the results are what we expect, we will reinforce that what was promised is coming true. People are much more likely to acknowledge value and continue with changes if the value is shown to them explicitly. For example, if senior leaders are consistently shown the value captured by project teams in applying Critical Chain scheduling, they will be far less likely to cut PMO funding. They will better understand the connections between continued funding and success.

If expectations are not being met, they may need to be reset. In that case, the implementation should be re-evaluated. It is important to fix problems early on. It is also important *not* to pretend that things are fine when they aren't. You won't be able to fool all of the people all of the time.

Player	Value	Sample Measurement
CEO	Improved predictability	On-time delivery
	Reduced cycle times	Standard cycle time vs. benchmarks, anecdotal evidence
	Increased efficiency	Number of projects completed vs. headcount
Project Manager	More credibility of schedules	Surveys ^a
	Better understanding of problems and their magnitude	Surveys
	Increased chance of hitting requirements dates	On-time delivery
Functional Manager	Better ability to predict and communicate resource requirements	Budget overruns/shortfalls, surveys
	Simpler assignment and management of tasks	Surveys
	Ability to say "no" or "not now" when appropriate	Interviews
Individual Contributor	Clear, stable priorities	Task progress statistics, ^b surveys
	Reduced chaos and multitasking	Surveys
Financial Officer	More reliable budgets	Deviations from plan
All	Consistent, disciplined communication	Checklists to determine process adherence

^aIt is relatively simple to construct short surveys, using (for example) a simple 1 to 5 scale, that measure how people feel about aspects of the value created through the scheduling process. Some types of information, for example schedule credibility, are difficult to get at in other ways. See also Newbold (2008, 136).

^bTasks should be worked in the priority indicated in the schedule, and the number of tasks active at any given time should not normally exceed the number of resources performing them. Checking these data in a computerized scheduling tool can give an indication of level of multitasking, and hence clarity of priorities.

TABLE 5-2 Critical Chain Benefits

The line from Validation to Urgency in [Fig. 5-4](#) indicates the ongoing need to understand and analyze the level of urgency. If people say that one thing is important (for example, cycle times) but behave as if another is important (for example, costs), we may need to re-think the implementation. If the implementation produces value and apparently reduces the level of urgency, we may need to

bolster the urgency with some additional actions, at least until the new processes are well established.

Many other cross-connections that are not represented here can occur during the CORE cycle. For example, expectations may need periodic adjustments based on the results of planning and implementation.

Root Cause	CORE Achievement
Lack of urgency	Urgency
Unwillingness to set clear expectations of value	Expectations
Lack of ownership in the solution	Commitment
Inadequate solution	Value
Inability to communicate value	Validation

TABLE 5-3 Mapping between Root Causes and CORE

As the cycle continues, expectations continue to be set and reset and commitment, value, and validation are built. All the elements may occur in parallel. Implementing and measuring, for example, do not normally stop as we communicate and replan. Some steps, such as replanning, will be skipped if they are not needed.

[Table 5-3](#) shows the direct relationship between the root causes from Figs. 5-1 to 5-3 and the CORE achievements in [Fig. 5-4](#). If CORE is implemented correctly and used as an ongoing process, it helps significantly in reducing or removing these root causes.

Simple Example: Cleaning the Room

CORE establishes trust that a set of changes will address an urgent need. It contains an implicit assumption: we wish the changes to continue into the future. We therefore put in place feedback loops to validate that continued trust is warranted. Let us consider a simple example.

Suppose you have a son, Billy, whom you want to clean his room. You want it done well, regularly, and without complaint. Some parents whine at their children until, perhaps, the child complies. Some threaten their children but when challenged fail to follow through on their threats. These approaches require little investment and may work a few times. However, they will ultimately fail because the child will realize that they have no reason to change. Consider instead the following approach, based on CORE.

1. **Urgency, vision, expectations:** Create a sense of urgency by explaining to Billy that he is not allowed to play after school until he has cleaned his room. Describe your vision of a “clean room.”
2. **Planning, commitment, implementation:** Work with Billy to plan how the “clean room” rule will affect his daily routine. You might make allowances for certain kinds of after-school events.
3. **Measure, validate:** Conduct inspections, explaining what he has done well and what he has not done well.
4. **Continue the cycle:** Allow Billy to play or require him to stay at home, depending on the results. Adjust the rules as necessary based on changing circumstances.

This approach is much more likely to cause Billy to gain trust that you mean what you say than threatening and complaining. Of course, it may also cause you, the parent, to reconsider your level of urgency. Consider that, if you leave out any of these steps, your chances of achieving the “clean room” vision will go down. Is that vision important enough to you that you will follow all the steps?

Simple Example: TOC Practitioners Group

Suppose you are interested in starting a TOC practitioners group to share best practices. You may have several reasons for this; for example, improving the level of implementation quality and thus the credibility of TOC in your area. How should you begin?

You will definitely need a target list of people who might take part. You will want to find out their level of urgency. What do they really care about? Where is their pain? If your group includes consultants, you may decide to increase their urgency by pointing out the advantages that will be gained by those consultants who attend.

From there, you will need a vision that ties the future to that urgency. Assuming the vision and expectations you set are sufficiently compelling, people will participate in the planning, further cementing their level of commitment. From there, you would need to continue to implement, measure, and communicate. If the practitioners group does not continue to provide value, participation will wane.

Other Processes

The feedback provided by CORE is essential to building trust in the urgency and consequences of change. In this section, I draw some comparisons between CORE and a few other well-known improvement processes. I encourage you to think about feedback (or the lack of it) in various processes with which you are familiar. For example, you might analyze which of the following processes contain feedback loops, and what kinds of changes those loops reinforce:

- Define-Measure-Analyze-Improve-Control^{[10](#)}
- Learn-Commit-Do^{[11](#)}
- Layers of Resistance^{[12](#)}

- Observe-Orient-Decide-Act¹³
- Ponzi schemes
- The scientific method
- TOC thinking process tools¹⁴
- TOC strategy and tactics trees¹⁵

CORE and Sales

Change requires sales, whether that means selling yourself on changes you need to make or selling others on changes they need to make.¹⁶ When I talk about sales, I don't mean the kinds of annoying tricks that are used by sales people to get you to part with your hardearned cash. Businesses that run on hard selling—pushing as hard as possible to get a sale—shouldn't expect a lot of repeat business. Buying is unpleasant and expectations are often far from reality. Instead, I am talking about selling that creates a win-win relationship between buyer and seller, a relationship that continues into the future. If a buyer of change is involved in such a relationship, she will continue the change. If she is not, she won't.

Solution Selling Step	CORE Element
Perform pre-call planning and research	Learn and analyze
Stimulate interest	Urgency
Define pain or critical business issues	Urgency
Diagnose and create a vision based on your solution	Describe a vision Expectations
Develop and manage evaluation plan	Plan and create ownership
Reach final agreement	Commitment
(Not included)	Implement Value
Measure success criteria	Measure results Validation
Leverage success	Communicate, re-evaluate, reinforce

TABLE 5-4 Solution Selling Steps and CORE Elements

CORE contains many elements of such a win-win selling process. It is closely related to a process called Solution Selling (Eades, 2004). Some important steps of the Solution Selling process are compared with CORE in [Table 5-4](#).¹⁷

There are a few interesting parallels and differences that you can see from [Table 5-4](#). The Solution Selling concept of “pain” corresponds to the CORE concept of urgency. In a selling situation, urgency is most commonly caused by pain. Consequently, Solution Sellers spend a great deal of effort understanding and exposing their buyers’ pain.

In order to describe the solution as it relates to the pain or urgency, both processes require communicating a vision. The vision connects the urgency to expectations of a future in which the pain is relieved.

Solution Selling is primarily targeted at reaching an initial commitment, while CORE is primarily targeted at creating and leveraging ongoing success. Some of the resulting differences are apparent from [Table 5-4](#). Solution Selling breaks the early stages, creating urgency and setting expectations, into more pieces. Those pieces are very important if you are driving toward an initial commitment, such as selling a senior manager on the idea of implementing Critical Chain in the first place. Very often, when selling efforts begin, people don’t understand their own urgency well; creating that realization requires work and effective tools. Urgency requires less emphasis if it is already well understood by the key players.

CORE, with its emphasis on ongoing success, places more weight on later steps like implementation and validation.

Plan-Do-Check-Act (PDCA)

Deming (1982) refers to the PDCA cycle as a helpful procedure to follow for improvement at any stage of

production.¹⁸ This cycle includes four steps: Plan (establish objectives for changes); Do (implement the changes); Check (measure the results); and Act (analyze the results).

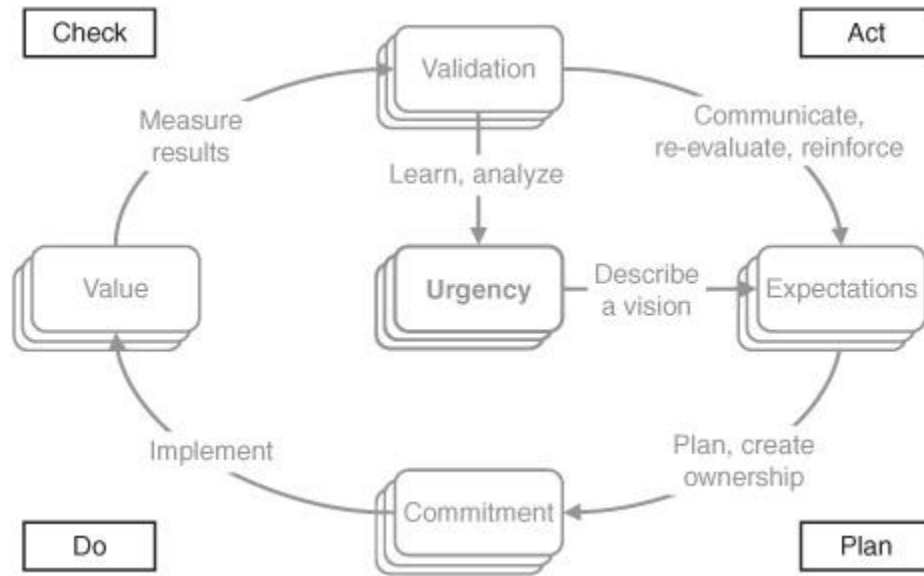


FIGURE 5-5 CORE and Plan-Do-Check-Act.

[Figure 5-5](#) shows how the PDCA cycle overlays CORE. The PDCA steps are analogous to the CORE actions. *Plan* corresponds to *planning and creating ownership*; *Do* corresponds to *implement*; *Check* corresponds to *measure results*; and *Act* corresponds to *communicate, re-evaluate, and reinforce*.

PDCA includes none of the CORE achievements (the rounded boxes). For the original purposes of PDCA, such as driving ongoing quality improvements, those elements may not be important. However, we have found all of them to be very important when people need to make long-term, system-wide changes.

CORE contains an unstated connection with PDCA that is important to understand. Deming suggests starting slowly in the “Do” step. As knowledge is acquired, the changes can be made more pervasive. We recommend a similar process when implementing Critical Chain in larger organizations:

start with a pilot in order to gain real-life understanding before making major changes to the organization.¹⁹

Five Focusing Steps

TOC practitioners often want to know how the TOC *Five Focusing Steps* (5FS)—*Identify, Exploit, Subordinate, Elevate, Go back to step 1*—relate to CORE.²⁰ The reinforcing loop in the 5FS process demonstrates the potential for constraints to move over time and the importance of dealing with those changes. It is a crucial loop; I have seen numerous examples of TOC production implementations that stagnated due to people's unwillingness to reidentify constraints and change behaviors as the constraints changed.

This points us to an important connection with CORE. The concept of *subordination* permeates the entire 5FS process. It means that everyone pitches in, working together synchronously to make sure that the focus remains on the constraints. In a sense, the 5FS are a guide showing what the people in the organization should subordinate to, namely the organization's goal and constraints. They can produce short-term benefits very quickly. CORE shows how to achieve that subordination, thus addressing the Uptake Problem—helping to cement the long-term benefits that come with ongoing improvement.

If people don't learn how to subordinate properly, an implementation of the 5FS may result in initial benefits, but the initiative probably won't last. I call this the "silver bullet" effect²¹: we are so tempted by the "silver bullet" of immediate benefits that we don't pay attention to negative branches shown in [Fig. 5-3](#). You need CORE to make the changes stick.

Implementation Planning

Thus far, I have described CORE and given simple examples of its application. However, a Critical Chain implementation is complex. It includes installation, training, business process changes, and new flows of information. It should result in significant changes to how people do their work, changes that must be synchronized across potentially dozens of functions and thousands of people. Much can and does go wrong. Ultimately, we want people to adapt our Critical Chain methodology to their environment in such a way that it becomes part of the organization's DNA.

Planning with the Cycle of Results

In order to use CORE to analyze an implementation plan, just follow the cycle. A number of questions will be immediately obvious, for example:

- What is driving urgency to change? Who experiences it? Who needs to experience it?
- Is there a vision that unifies the urgency experienced by the different players?
- Have expectations been set? For whom? Who will set their own expectations, and what is the impact of that?
- Has the planning accomplished buy-in?
- Who is truly committed? In other words, whom can you trust to take the lead?
- Where do you expect value to be created?
- What measurements will be used to validate that value was created?
- How is that value going to be used, both to (re) shape expectations and to adapt the implementation plan?

If we apply this approach to the steps in [Table 5-1](#), we may discover a number of missing elements, without which we will not get the CORE achievements (the rounded boxes). Adding these elements allows us to hope that we will overcome the root causes shown in [Table 5-3](#) so that the implementation plan will sustain itself long-term. Here are a few things to consider when applying CORE to [Table 5-1](#).

Urgency

A basic process for a group to raise performance to a new level was laid out many years ago by the pioneering social psychologist Kurt Lewin: unfreeze the present level, move to the new level, and freeze at the new level.²² Unfreezing can most easily happen through a sense of urgency, which is why urgency is so crucial.²³

Do we truly understand the level of urgency that different people are experiencing? Is it adequate to “unfreeze” people’s behaviors? We may have a great Critical Chain champion in an organization, but if she is the only one with a sense of urgency then she will have a difficult struggle. I have seen many champions ultimately lose heart because they did not back up their passion with the ability to generate and communicate a sense of urgency that resonated with their audiences.

Tip: Do the research and find the urgency. The research usually requires interviews, with the questions targeted toward understanding the individuals’ personal sense of urgency.

Expectations

People commonly communicate expectations through a vision. The vision is important, but it is seldom enough.

Given that different people will have different roles and expectations, we have found that a *Communication Plan* is usually needed. The Communication Plan is typically a spreadsheet that helps keep track of who is communicating what to whom, including:

- Expectations of different stakeholders and how expectations have been set.
- Marketing to groups not directly involved with the implementation.
- Feedback between the PMO, Steering Team, Implementation Team, and others.

Tip: Maintain a communication plan.

Commitment

Who should be involved with the planning? I have already mentioned the Steering Team and Implementation Team concepts. These groups need to leave their stamp on the plan. In a complex implementation involving many people, there will be many levels of planning to coordinate. Always remember that an important purpose for planning is to allow people to develop ownership.

Tip: Use planning to build ownership.

Value

It seems obvious that any change initiative should create value. However, strangely, we often find that people have not fully thought through answers to questions like, “What value?” and “For whom?” Companies invest millions in

Product Lifecycle Management (PLM) and Enterprise Project Management (EPM) systems without a clear idea of how those systems will benefit the organization or the individuals in them. Mismatched expectations among buyers and sellers can cause EPM implementations to drag on for years.

Tip: Identify the expected value and how it will be achieved and measured. Start collecting data early; there is no reason to wait.

Validation

Often we find that people assume an implementation will continue to thrive and produce benefits once it is well begun. It is true that benefits gained early can help to justify and give momentum to the implementation. Unfortunately, because the benefits of Critical Chain start well before the organizational DNA has changed, an associated “silver bullet” effect can lead to the negative branches shown in [Fig. 5-3](#).

Tip: Continue to collect and analyze implementation measurements, such as those in [Table 5-2](#), so that your implementation will continue to adapt and improve.

Traps

A number of conceptual traps lay waiting in implementations, traps that people fall into without thinking. You should review these traps periodically, just to make sure you have not fallen into one of them.

It's Not about You

We have a tendency to believe that our own opinions and actions are more important than those of others. We look at

what we need to do and what we need to get other people to do, without considering what they need to get themselves to do. We sometimes forget that others may have valid ideas as well.

Instead, think of an implementation as moving from “I” to “They.” This might be a progression, for example, from a world in which you as a facilitator take maximum responsibility for the implementation, to a world in which it would proceed even if you were run over by a bus. It goes like this:

I: Where we have to start

We: Better

They: Best

We make ourselves obsolete, bridging the gap between “I” and “They,” by using CORE concepts: setting expectations, building ownership and commitment, and creating and communicating value. That way we build ownership in the people who will eventually have to take responsibility.

Tip: Ask yourself: am I taking on too much? Am I delegating enough? Are the right people taking ownership?

Broken Trust

Have you ever heard a management team say, “If we implement the following technology, we’ll get the following incredible benefits,” only to find that after months or years of hard work the implementation fails to produce anything close to those benefits? In my experience, this is common for improvement initiatives; as shown in box 16 of [Fig. 5-3](#), it is not common practice to set clear, realistic expectations for an initiative. This is also a perfect example of broken trust. The promise of unrealistic expectations was made and

ultimately broken. Even worse, bad news travels quickly. If you break trust with one group, you have to believe that many others will hear about it. It is no wonder we so often find people who have little faith in their organization's ability to change.

CORE should be used to build and retain trust. We set realistic expectations, take actions to achieve those expectations, and visibly confirm that we have met or fallen below expectations. Either way, we continue learning. This works from an important principle: *The easiest way to regain broken trust is never to lose it in the first place.* If you make a habit of setting unrealistic expectations, you will often be disappointed. But even more important, you will build a culture of mistrust.

Before beginning any initiative, think through the realistic expectations that you are going to set with different stakeholders. The expectations should be adequate to address the vision and associated urgency. Communicate them broadly, with a Communication Plan and marketing campaign, so that expectations are not at the mercy of the grapevine. Then, when the initiative is underway, communicate how it is going and why.

Tip: Set realistic expectations and communicate progress frequently.

What is "Done"?

Very often, during the Critical Chain scheduling process, we find that people don't know what "done" means for certain tasks or projects. Often someone will understand a task in general terms, but be unable to say at what point it should be handed off. That can lead to quality problems when work is handed off before it is ready, or excess time and effort when work continues beyond what is needed. Team

members should always try to clarify the meaning of “done.”

During implementation planning, the reverse problem occurs. We create various plans and tools that have implementation work defined as discrete tasks. That may be extremely valuable, but if we are not careful—if we focus too much on things that can be declared “done”—we can lose track of the “level-of-effort” work that needs to be performed steadily into the future, such as:

- Communication planning
- Measurements of quality and performance
- Mentoring for behavioral changes, such as prioritization and reduced multitasking
- Transfer of methodology elements to the people doing the work
- Process ownership and improvement

These represent work for which it is difficult and often dangerous to declare “done.” Still, we have often seen exactly this occur: people assume that these kinds of activities are discrete tasks. They decide, for example, that implementation planning or measurements are no longer necessary once the implementation seems to be going well. They declare “done” and eventually experience the problems described in [Fig. 5-3](#).

Not all implementation work fits neatly into a project plan. Put another way: *If all your implementation work fits neatly into a project plan, you are missing something.* This should not be surprising, because the Project Management Institute describes a project as a “temporary endeavor” (2008, 434), and we want an implementation to be an ongoing process.

We can apply this principle immediately to the root causes as we are tempted to convert them to tasks. Do we create a sense of urgency and then declare “done,” or do we continue to communicate and reinforce the vision and urgency? Do we plan how to overcome the initial obstacles, or continue to evaluate new ones? Do we declare “done” for an implementation, or assume that it is part of a process that will never be “done”? We cannot answer these questions by adding tasks to an implementation plan. We need measurements of value and measurements of implementation status, as shown in [Table 5-2](#). We need communication of expectations and results. We need to understand urgency and feed it into daily and weekly communications. And we need to have people who are responsible for filling these needs.

Maintain a Communication Plan. Create oversight processes to make sure quality problems are addressed. Hold regular Steering and Implementation Team meetings in which critical issues are discussed. Create forums for internal experts to share knowledge. More broadly, create a CORE culture that rewards people for communicating honest expectations and results, whether those results are deemed good or bad.

Tip: Plan for work that will never be “done.”

Summary

Real organizational change does not just mean admitting that things need to change or trying different things. It means actually changing the habits that govern how people work—changing the organizational DNA. Improvement initiatives that require real change have, at best, a mediocre record of producing and sustaining long-term benefits, and this Uptake Problem is pervasive. Root causes for it include:

- Lack of urgency
- An inadequate solution
- Lack of ownership in the solution
- Unwillingness or inability to set clear expectations of value
- Inability to communicate value

These root causes can be addressed by applying CORE, a process developed by Pro-Chain Solutions over the course of many years facilitating Critical Chain implementations.

CORE requires the following steps:

1. Learn and analyze to find or create the shared *Urgency*.
2. Define and communicate *Expectations* using a common vision.
3. Build *Commitment* through planning.
4. Create *Value* through the implementation.
5. *Validate* the results through measurements.
6. Continue the cycle into the future.

CORE is a process for selling change that addresses the root causes by building trust in the change initiative over time. The ProChain experience indicates that the combination of CORE and best-in-class solution components results in successful and long-lasting implementations.

The CORE feedback cycle can be applied to simple and complex situations. Some associated traps can be avoided by transferring ownership of the implementation, creating

realistic expectations, and acknowledging the fact that some implementation tasks will never be completely “done.”

Ability to change quickly—“agility”—can be a tremendous competitive advantage. For an organization to be truly agile, it must be able to respond rapidly to changes in markets and technologies. In order to make appropriate changes that stick, its people must have the patience, discipline, and flexibility to build trust in those changes. CORE is an important tool to build that trust.

References

Cialdini, R. B. 1993. *Influence: The Psychology of Persuasion*. New York: William Morrow and Company.

Covey, S. R. 1989. *The Seven Habits of Highly Effective People*. New York: Simon and Schuster.

Deming, W. E. 1982. *Out of the Crisis*. Cambridge: MIT CAES.

Eades, K. M. 2004. *The New Solution Selling*. New York: McGraw Hill.

Dettmer, H. W. 2007. *The Logical Thinking Process: A Systems Approach to Complex Problem Solving*. Milwaukee, WI: ASQ Quality Press.

Duck, J. D. 2001. *The Change Monster*. New York: Three Rivers Press.

Goldratt, E. M. 1990. *What Is This Thing Called Theory Of Constraints, and How Should It Be Implemented?* Great Barrington, MA: North River Press.

Goldratt, E. M., Goldratt, R., and Abramov, E. 2002. Strategy and tactics. <http://www.vancouver.wsu.edu/fac/holt/em534/Goldratt/Strategic-Tactic.html>.

Gupta, S. 2005. Critical Chain: Successes, failures, and lessons learned. Presentation at the 3rd Annual TOCICO Conference, November 2005, Barcelona, Spain.

Hobbs, B. and Aubry, M. 2006. Identifying the structure that underlies the extreme variability found among PMOs. Newtown Square, PA: Project Management Institute Research Conference.

Koenigsaecker, G. 2009. *Leading the Lean Enterprise Transformation*. New York: Productivity Press.

Kotter, J. P. 1996. *Leading Change*. Boston: Harvard Business School Press.

Kotter, J. P. 2008. *A Sense of Urgency*. Boston: Harvard Business School Press.

Lean Enterprise Institute. 2008. Backsliding is back as the biggest obstacle to lean transformations. http://www.lean.org/WhoWeAre/NewsArticleDocuments/Obstacles_adden_dum_release08.pdf.

Lewin, K. 1997. *Resolving Social Conflicts and Field Theory in Social Science*. Washington, DC: American Psychological Association.

Newbold, R. C. 1998. *Project Management in the Fast Lane: Applying the Theory of Constraints*. Boca Raton, FL: St. Lucie Press.

Newbold, R. C. 2008. *The Billion Dollar Solution: Secrets of ProChain Project Management*. Lake Ridge, VA: ProChain Press.

Project Management Institute. 2008. *A Guide to the Project Management Body of Knowledge*. 4th ed. Newtown Square, PA: Project Management Institute.

Retief, F. 2009. Critical Chain vs. Pooled risk scheduling. See <http://www.mpsys.com.au/downloads> for download information.

Richards, C. 2004. *Certain to Win: The Strategy of John Boyd, Applied to Business*. Philadelphia: XLibris Corporation.

Scheinkopf, L. 2000. *Thinking for a Change*. Boca Raton, FL: St. Lucie Press.

Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>.

About the Author

Robert C. Newbold, CEO and founder of ProChain Solutions, is one of the world's leading experts on project scheduling and management using the Critical Chain approach. Rob is a frequent writer and speaker on the subject of project management. Over the past 25 years he has developed process improvements in the fields of health care, manufacturing, and project management. He is the author of *The Billion Dollar Solution* (2008) from ProChain Press and *Project Management in the Fast Lane* (1998) from St. Lucie Press, and holds degrees from Stanford University, State University of New York (SUNY), Stony Brook, and Yale University.

CHAPTER 6

Project Management in a Lean World —Translating Lean Six Sigma (LSS) into the Project Environment

AGI-Goldratt Institute

Introduction: It's a Lean World

For most large organizations in the Western Hemisphere, the call to pursue a discipline of improvement began with the 1980s NBC broadcast of “If Japan can . . . Why can't we?” Many embarked on the quality movement of putting human and financial resources toward that commitment. Investing in training from Dr. Edward W. Deming, Dr. Taiichi Ohno, and Shingeo Shingo as well as juggling the onslaught of new training and consulting organizations that emerged, the mid to late 1980s saw the introduction of a myriad of techniques—most seeming to have a three-letter acronym. Whether it was SPC (Statistical Process Control), TPS (Toyota Production System), SMED (Single Method Exchange of Die), JIT (Just-in-Time), or TPM (Total Productive Maintenance), external and internal experts with different techniques descended upon the business units to form numerous Process Improvement Teams, all competing for the same resources that were already fully needed just to run the business.

Motorola is credited with the invention of the Six Sigma methodology. Those inside Motorola saw the power of the various techniques from TQM, Deming, Juran, and others

and evolved them to a management system that was focused on improvement and the bottom line. First aimed at processes within manufacturing, Motorola then developed the elements to embed it within their operating culture.

Thanks to James Womack and Daniel Jones through their book, *Lean Thinking* (1996), the tools of the quality movement now had a framework to work more collectively—the *Lean Principles*. *The principles of specifying value and the value stream, creating smooth flow, and enabling the customer to pull value and the pursuit of perfection ensured the process of improvement would be ongoing.* (For a good synopsis of Lean and Six Sigma methodologies, please refer to [Chapter 36](#), “Combining Lean, Six Sigma, and the Theory of Constraints to Achieve Breakthrough Performance by AGI-Goldratt Institute”).

Both Lean and Six Sigma continue to be heavily embraced by the private and public sectors and have become more and more integrated as Lean Six Sigma (LSS). Both are well developed. Both enjoy the support of many top executives, line managers, and vast numbers of employees who have been trained to one degree or another in these disciplines. Let’s face it, for most of us it is a Lean world!

What implication does this have on the project environment? The attention on Lean Six Sigma continues to grow. There are whole offices and departments set up for LSS. Funding availability seems plentiful in relationship to other needs. There are growing numbers of experts in LSS from white belts to green belts to black belts. These many experts and efforts all result in a broadening of the application of LSS from the shop floor to the whole organization—including the project environment!

What Is the Project Environment’s Point of View to Being Leaned?

As the LSS efforts broadened into the project environment, there was less than an enthusiastic greeting. Most project managers and resource managers felt that they were already working in a pretty lean world—lean on resources, lean on time, and lean on funding. Many project managers felt that they were already asked to do the near impossible—sit on top of an elephant balancing on a ball on a high wire 20 feet in the air without a net ([Fig. 6-1](#)).

It feels pretty lean when one feels they are already working
without a net!



FIGURE 6-1 PM’s point of view. (©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.)

In trying to “lean” the project environment, there have been a few seemingly insurmountable obstacles. To begin with, like supply chain environments, project environments are made up of a system of systems. This increases the difficulty of deciding not only where to focus but also how to determine the most opportune areas of waste and value. Additionally, when applying definitions and techniques for improving the areas of productivity, focus, value, waste, and

variation to a project-based system, there appear to be disconnects as LSS's techniques and definitions were developed for the manufacturing environment and appeared to not readily apply to the project environment without significant translation. Couple that with the fact that traditional project management techniques contained in the *project management body of knowledge* (PMBOK) have not necessarily integrated Lean. No wonder there has been a lukewarm if not cool reception. Let us look at these issues more thoroughly one at a time.

Project Environment System of Systems

There are four systems within a multi-project environment. They are the task management system, the individual project system, the portfolio of projects system, and the resource management system.

The *task management system* ([Fig. 6-2](#)) consists of the list of tasks or group of interrelated tasks where a person is responsible for ensuring that all the elements for that task are completed by the scheduled date (and often within the cost estimated).

The detail under the "task" does not generally show up in the project schedule, only the overall task. If one were building a house, this task might be called "complete electrical wiring." The crew chief would have one electrical crew to oversee pulling 110-volt wiring to lights and outlets; another perhaps running 220-volt wiring for some appliances; and another setting up the electrical panel.

System of Systems Project Environment

In many project environments, there is schedule and/or cost management at task or group of tasks level.

Task listed within the project:

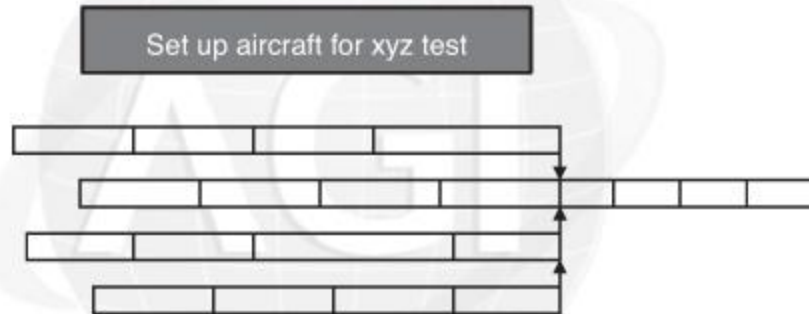


FIGURE 6-2 Task management system. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

System of Systems

At the project level, we have the project system—the sequence of tasks, handoffs, and deliverables that when accomplished deliver the desired outcome.

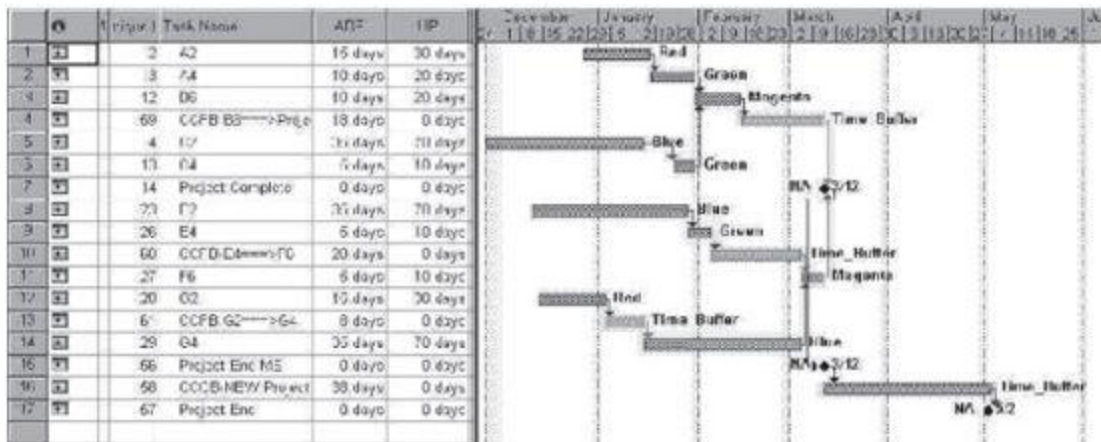


FIGURE 6-3 Project environment. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

The *individual project system* consists of the sequence of tasks, handoffs, and deliverables that when accomplished deliver the desired outcome. The individual project system must manage the delivery of content within a committed time and budget. Very often, scheduling begins with the various resource functions listing their tasks and time (or level of effort) as stand-alone elements ([Fig. 6-3](#)).

The individual project content commitments are made independently of other projects' task work for shared resources. Even when shared resourcing is considered, little notice is given to the impact of variability on the releasing of a resource from one task to another. In addition, where project-to-project dependency exists, often project commitments are made without consideration of the impact of variability of one project on another. An example might be when the organization is developing a project where the output would be used by another or several other projects, such as the development of a new microprocessor that will be utilized in each successive product platform.

At the *portfolio of projects level system*, all the projects are grouped by product type, business type, or organization type and must be managed to ensure that each customer is satisfied. Unfortunately, the need dates of the customers are independent and are not necessarily able to be coordinated across a portfolio ([Fig. 6-4](#)).

At this level, conflicts between projects for limited shared resources become more visible. Unfortunately, there are often compromises made—which projects will be given higher priority for resources versus others, and many projects struggle as they have to manage without the benefit of being the “hot” project.

Finally, at the *resource management level*, the organization needs not only to plan what capacities they must have to support current and future project work, but also handle how to deploy the current resources to the queue of the tasks for each project—each with a project

and/or portfolio priority. The managers of this system constantly juggle the capacity available and task execution priorities (Fig. 6-5).

The resource manager is often put into the position of switching resources back and forth to the new squeakiest wheel (task), trying to spread the capacity where it might do the most good against a seemingly never-ending queue.

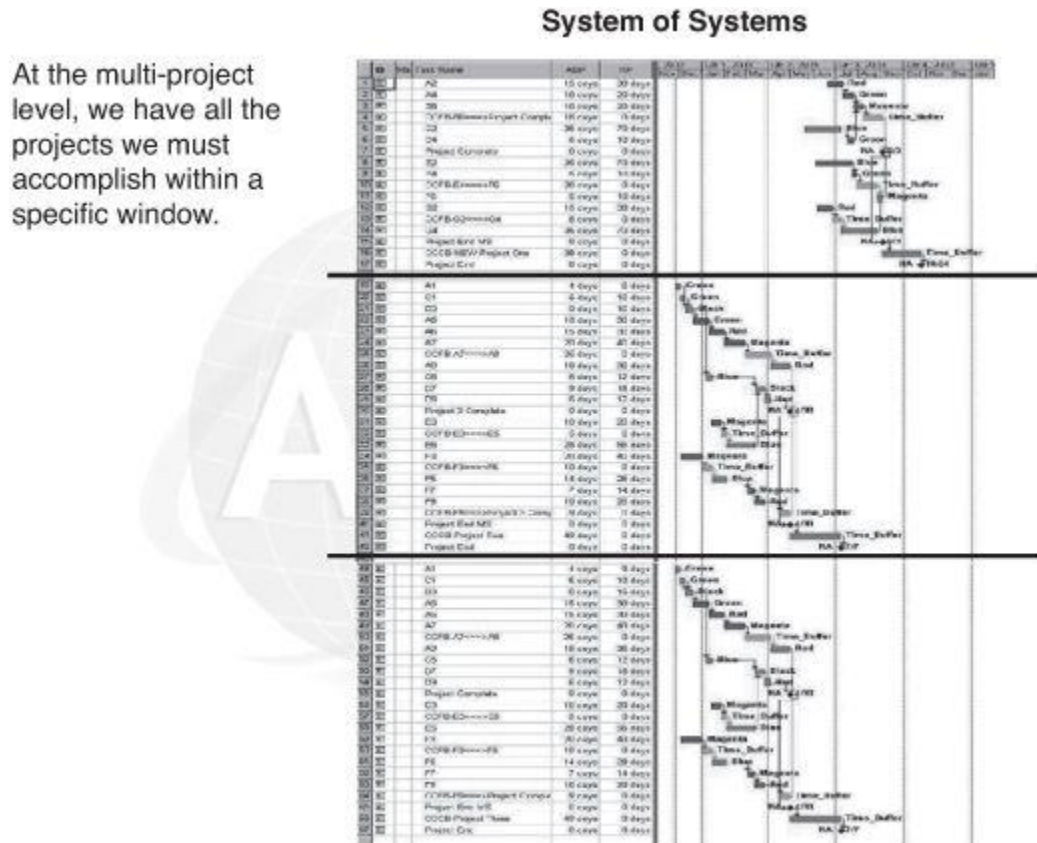


FIGURE 6-4 Multi-project environment. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

What Do We Improve?

With all these different systems and owners, it appears that approaching system improvement in project environments is

like the “The Blind Men and the Elephant,” the Indian fable immortalized in the poem by John Godfrey Saxe (1873, 77-78) Project management system improvement has some interesting challenges. There are many owners of these different “systems,” each with their own view of what needs to improve. As long as these systems are not aligned to work in concert, there will be little opportunity for real improvement. This means that the relationships between these systems need to be understood. Ultimately, the capacity of the organization (either based on its limited capacity resources or the amount of a type of work that be can be taken on in a window of time) should dictate how much work is accepted in the portfolio or pipeline. Only then can individual project commitments be made. Task priorities then should be based on this release of work and the actual availability of “ready to work” tasks. Adjustments should only be made to task list priorities when there is objective data that the project require for the task to be expedited. The key to improvement is this alignment of these systems of systems and, with this understanding, translating Lean Six Sigma to drive value, minimizes waste.

Translating Lean into the Project System of Systems for Improvement

Lean manufacturing could be summarized by what has been attributed to Eiji Toyoda in describing a pillar of the Toyota Production System: “providing *exactly* what the customer wants; *when* the customer needs it; in the *correct* quantity and in the expected *sequence, without defects*; at the *lowest possible cost*.” We must consider the importance of this concept, but apply it to each of the system of systems in a project environment in a way that aligns the system.

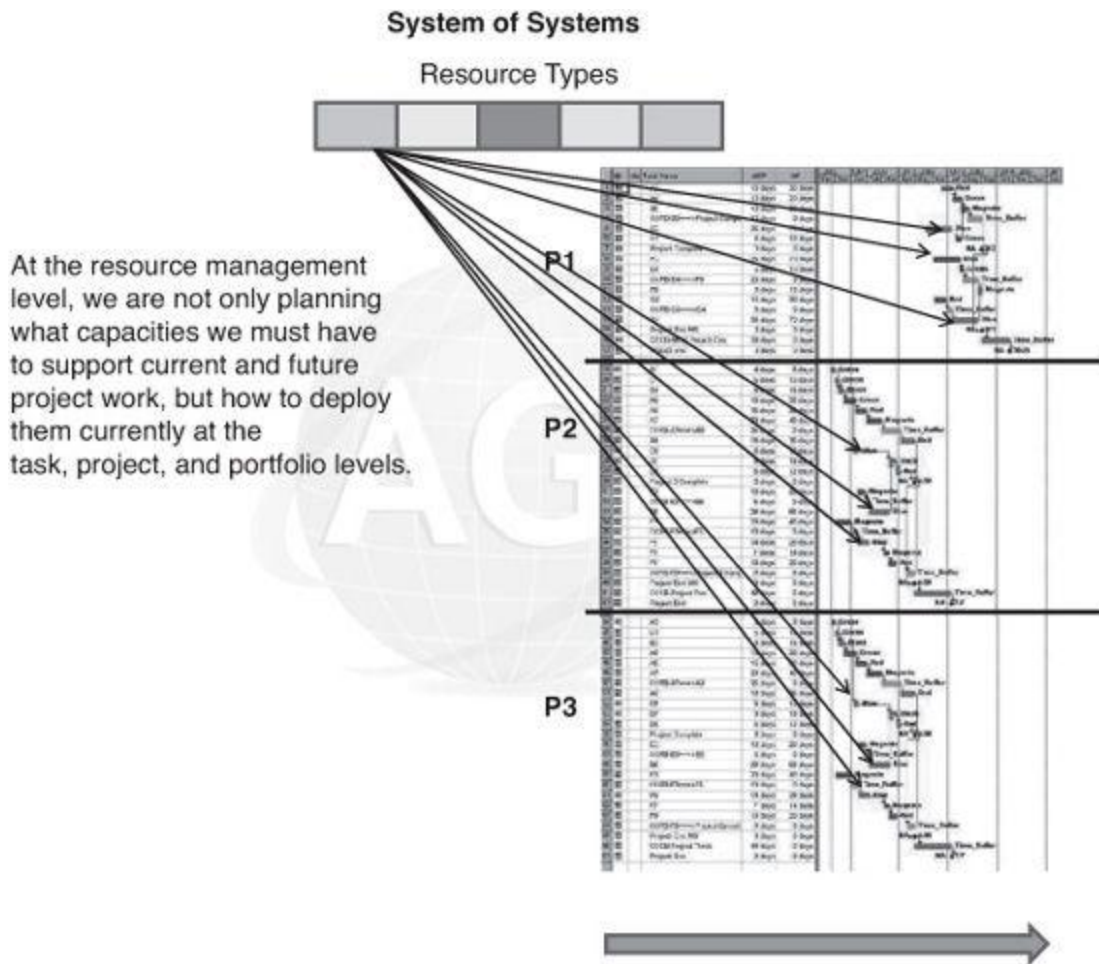


FIGURE 6-5 Multi-project resource management. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

In a multi-project environment, we start by aligning the system of systems with the *capacity* of the organization and the *portfolio* of work. Lean would mean taking on the right quantity of projects, based on the organization's capacity to do work (within a window of time), with the correct content, as quickly as possible to meet each project's needed commitment date. For those projects that are agreed to be taken on by the portfolio, Lean would mean accomplishing the right tasks, in the right sequence, with the correct quality, as quickly as possible to deliver exactly what the customer wants, when the customer needs it. From there, Lean as applied to the task priorities would translate as

having the right tasks assigned, in the right sequence, utilizing the correct resources. Next, Lean Task Management would mean ensuring that the right tasks are executed, at the right time, delivering the correct content with the correct quality, as quickly as possible (Fig. 6-6).

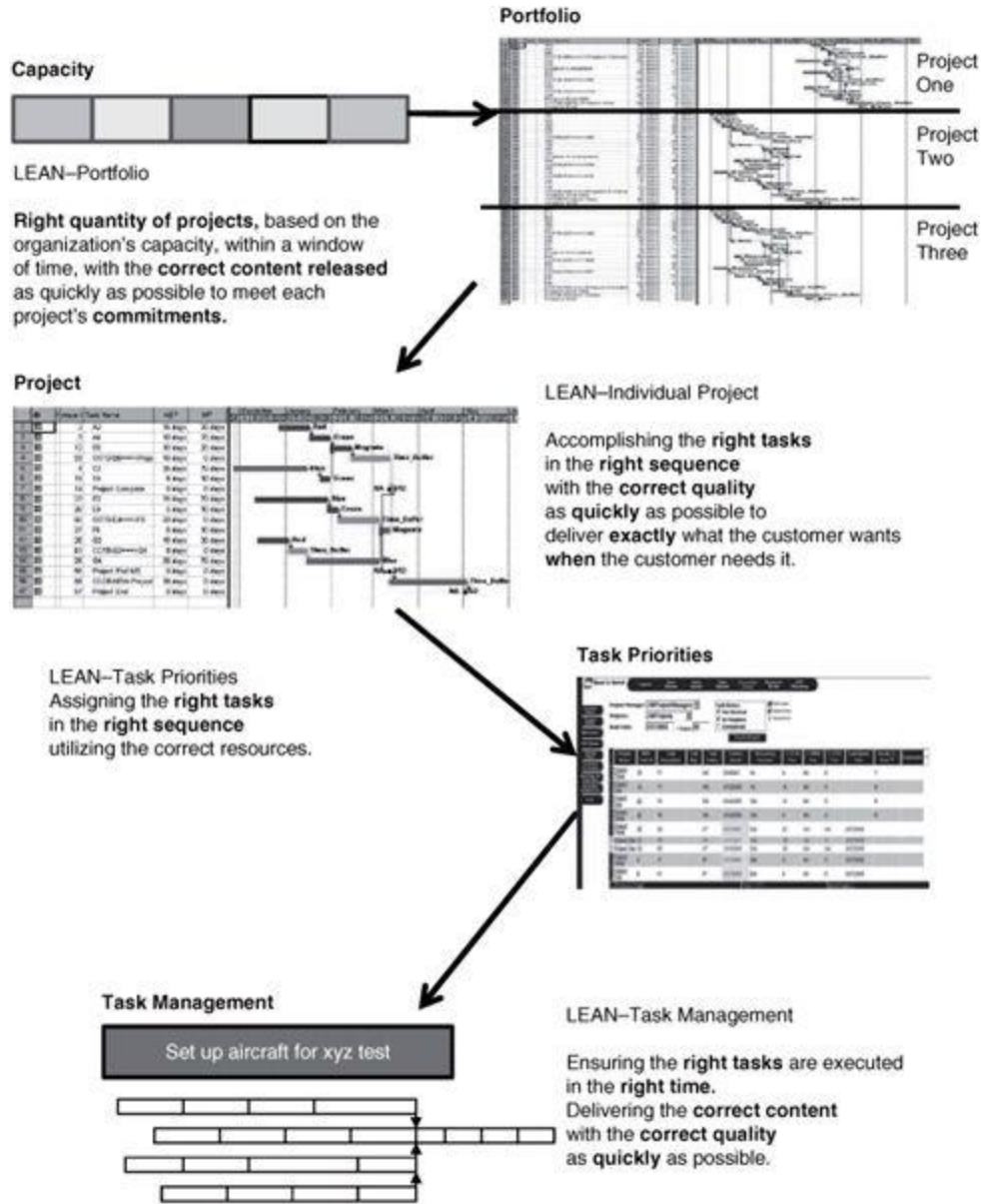


FIGURE 6-6 Aligning the systems in a project environment. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

Addressing the Disconnects in Lean Techniques for Project Environments

As stated earlier, there are obstacles in applying LSS to the project environment. We have already addressed the issue of the system of systems nature of the project environment. It is now time to turn our focus to those disconnects with applying definitions and techniques derived from a manufacturing environment and applying them directly to a project environment. In particular, we will look at what is needed to improve productivity, focus, and value, and to eliminate waste and variation.

What is productivity in a project environment? One might be tempted to look at the percent load on the various resources versus their availability in deciding if the project environment is more or less productive—after all, this is where the project organization's costs and investments are. However, this would be taking the traditional efficiency concept from the manufacturing floor and directly applying it to the project resources. The organization would only be measuring how active their resources were rather than how productive they were. Consider this: working out on a treadmill generates a lot of sweat and does provide a cardiovascular workout. Yet, if your goal is to go from Point A to Point B, nothing has been accomplished—there is activity, but one is not productive in getting to Point B. If our goal is to go from Point A to Point B as quickly as possible, then running faster from Point A to Point B is more productive than running slower or stopping periodically to go shopping, eat, or do email. A project's throughput is only achieved when it is complete. How quickly an organization can sequence in that project to achieve throughput is based on the organization's capacity in a window of time and is driven by how much work the resources can accomplish. It would follow that speed of execution of the right tasks

accomplished with the correct content and quality drives speed of execution of each project and our capacity for the pipeline of work. Productivity must be viewed from the task perspective—the speed to accomplish the task.

Are we driving the productivity of tasks? Are the metrics within the project environment driving in productivity or do they actually drive in waste? In some organizations, some key metrics are items such as hours charged out per person, resource utilization, and earned hours. These metrics have little or no relationship to whether the hours worked were on the right tasks. In looking at an example from Earned Value (EV), we have two environments ([Fig. 6-7](#)).

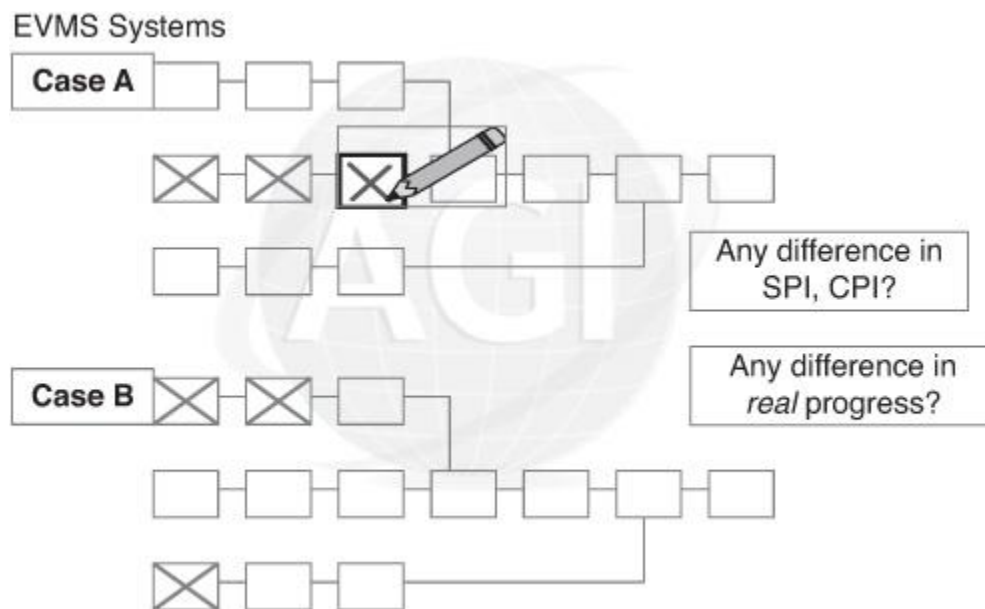


FIGURE 6-7 Activity versus productivity. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

The top one shows the case when we earn hours on the longest pathway. The second shows that the same number of hours has been earned, but the tasks that drive the project schedule have not been touched. The metric of earned hours and subsequent indicators of cost and schedule performance (SPI [Schedule Performance Indicator]

and CPI [Cost Performance Indicator]) may not alert the organization that they are not being productive on the tasks that drive project completion and the achievement of throughput.

How does the project environment use the five areas identified by Womack and Jones (1996) as the key principles of Lean to ensure improvement would be ongoing? The five are:

1. Specify value from the standpoint of the end customer.
2. Identify all the steps in the value stream.
3. Make the value-creating steps flow toward the customer.
4. Let customers pull value from the next upstream activity.
5. Pursue perfection.

The Five Principles of Lean Applied to the Project Environment

Specifying Value

How do we specify value in projects? Lean principles start with an attempt to *define value* in terms of *specific products with specific capabilities offered at specific prices through a dialogue with customers*. Taking the time to define the project value will alleviate some common problems found in the project environment, such as project definition being too vague, lack of stakeholder support/participation, scheduling

without really knowing the true scope, and scope creep. A simple technique for specifying value revolves around answering some key questions. Who is the customer of this project? Is there more than one? Is our own organization also requiring value from this project? Is the objective and scope sufficient to solve each of the project's customers' problems so that we will not have to expand or redo the project? This means we must first establish the problem statements for each customer and only then specify what we must deliver to solve that problem and create value.

Identify Steps in the Value Stream

Once we have defined the value from the standpoint of the end customer, we *must identify all the steps in the value stream*: the project structure of arrows, tasks, and resources that create that value. To ensure each task, relationship, and resource is not wasted, we can use some guiding questions. Is each task and path dependency necessary to achieve the customer objective? Is it creating value? If a task is not creating value, is it necessary for satisfying a boundary condition of the project (e.g., may not use outside contractors on constructing competition-sensitive operating equipment)? Does this task meet the correct exit criteria to provide the correct input for its successor task?

Make Value-Creating Steps Flow towards the Customer

As we plan the project, we need to ensure that the *value-creating steps flow towards the customer* and that the project deliverables solve each customer's problems. We should ask whether this task dependency is necessary to ensure we do it right the first time (or to minimize iteration variability) and is it worth the time investment of waiting for the predecessor task to complete? Is the investment of this

type of resource (high-level skill) in this task appropriate for the investment of the loss of the critical resource being tied up? As the value stream is mapped, what we call the project network, hopefully most of the steps will be found to create value. Additional steps may be listed and not add value to the product or service. Those steps that create no value and that should be eliminated are called *Muda* or Waste.

How do we decide if we have the correct tasks and the correct dependencies between tasks? The answer can be summarized as the correct tasks and arrow dependencies are those that are necessary to deliver the project scope and support their successor tasks to enable speed and quality. Do we have all the tasks that create value? It is important in projects not only to ensure that we only have tasks and arrows that are needed to create value, but also that we have no omissions of tasks that are needed to deliver full value.

Let Customers Pull Value from the Next Upstream Activity

In executing projects, we must execute in a way that lets each task's customer (successor task, deliverable) *pull value from the previous upstream activity*. As a project schedule is followed, it must be followed to ensure task execution, arrow dependencies, and resources assignments occur as planned to minimize waste and create value for the customer.

Efforts to improve the project system of systems must address the waste that slows task accomplishment, wastes our limited resources 'time, and increases the costs in projects. Waste comes from two main areas. The first area of waste can occur during planning; whether through identifying the wrong tasks or arrow dependencies; incorrect assigning of resources to a task; missing tasks, or incorrect or incomplete customer requirements. The second area of

waste occurs during the execution of the project from the misalignment of priorities, misuse of limited resources, or misaligned behaviors. We address waste issues within the project plan during project planning and scheduling. We address waste in project execution with the alignment of the system of systems.

What is waste in a project environment and will I know it when I see it? Dr. Taiichi Ohno (1988) identified seven categories of waste (to which an eighth category has recently been added). Many of the definitions for these categories are manufacturing based and not project based—yet the categories are very powerful to drive out waste, create speed, and increase capacity in the project environment. These categories are translated as follows.

Categories of Waste in a Project Environment

The first category of waste is *overproduction*. In the project environment, this can translate into starting a path or task before it is available to start or assigning resources to any task because you have the resources and not because there is a task needing that resource or that quantity of resources. Additionally, overproduction might be seen as doing a task as part of the project, when in fact it is not part of delivering the value of the project.

[Figure 6-8](#) depicts an example of what was planned versus how it was executed. The organization ends up spending additional time on a task, more than was needed, and tying up resources longer for no additional value or speed.

The second category of waste is *waiting*. Since productivity should be defined as how fast we complete a task and hand it off, then when a task is interrupted and waits for a resource that is pulled away to work on other tasks at the same time, the task experiences waste in the time it waits or is idle while the resource works on another

task. This is often the case when a resource is multitasked ([Fig. 6-9](#)).

Another example of waiting occurs when a predecessor task completes its work, but does not pass on that work to the successor task. The successor task experiences waste by waiting for its handoff.

The third category of waste is *transportation*. Transportation waste in projects occurs when an incorrect predecessor-successor task dependency is identified, resulting in an unnecessary delay waiting for a predecessor task to be completed for an input that is not necessary for the successor task to start. Another example is when a review that generates a “looping” back or rework loop is later in the process than it should be, lengthening the project’s overall time by the time it takes to redo the earlier tasks.

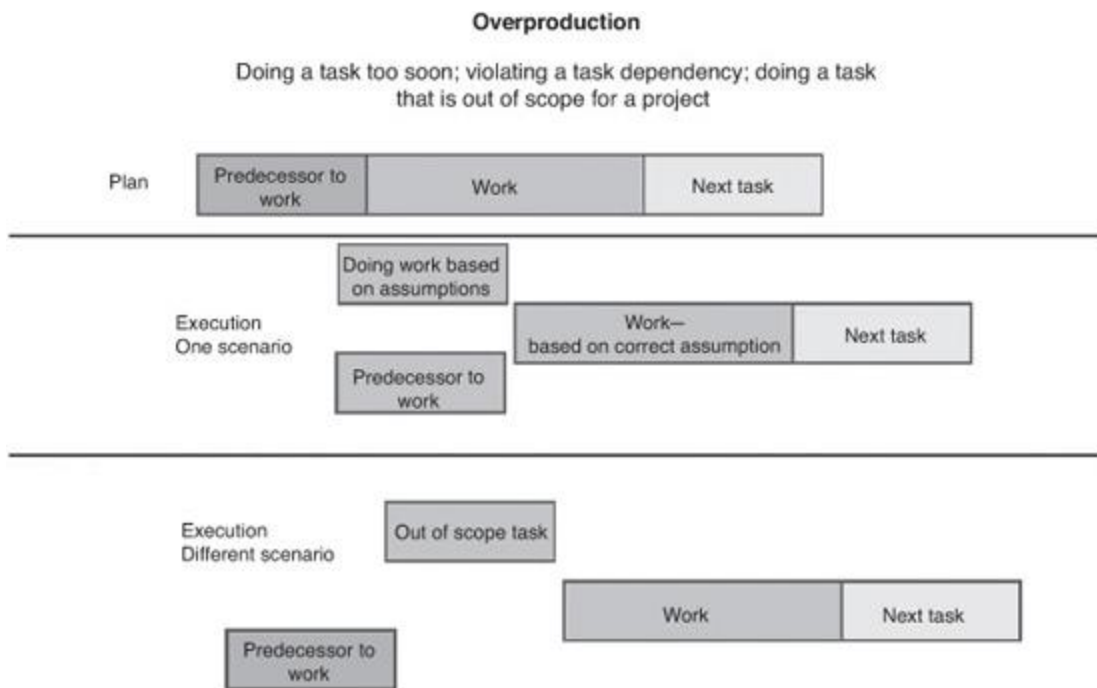


FIGURE 6-8 Overproduction. ©1991–2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

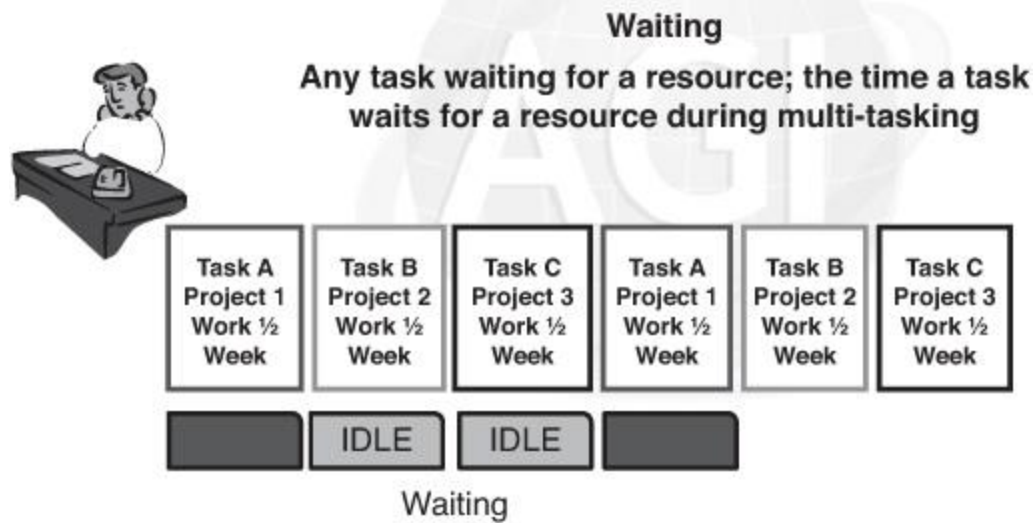


FIGURE 6-9 Waiting during multitasking. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

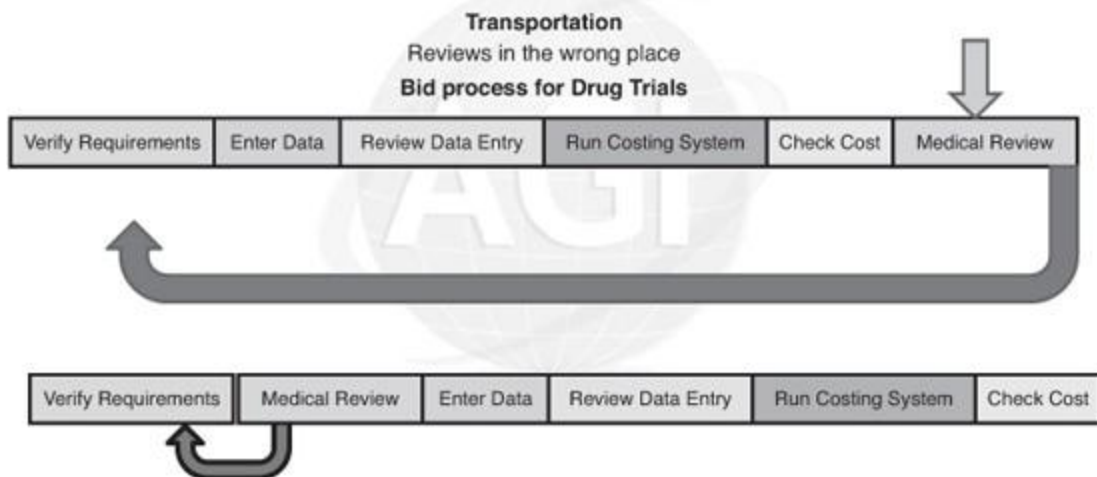


FIGURE 6-10 Transportation: Reviews in the wrong place. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

[Figure 6-10](#) shows a medical review of internal requirements needed to meet customer requirements for a

specific drug trial occurring after the time-intensive costing process. This review could be done early in the process prior to the more time-intensive tasks, shortening the quantity and time investment of tasks that might need to be reworked.

The fourth category of waste is *excess inventory*. In a project environment, excess inventory is represented by elements of too much task work in progress, or resource/resource groups accomplishing more tasks than the organization can process. Additionally, some projects require too many supplies, unneeded files, or unnecessary copies of documents or prototypes. Excess inventory also occurs when we require more of a skilled or limited resource than the task requires. In some project environments, the project dedicates resources to the project for its entire length.

[Figure 6-11](#) demonstrates the amount of a particular resource's time budgeted to the project versus the actual need for that resource, creating an inventory of available hours that will be used by the project but not necessarily drive value.

The fifth category of waste is *excess motion*. Excess motion occurs in projects when time is taken on a task that is not inherently needed to accomplish the task to create value. Holding onto a task that is complete and continuing to polish the output or searching for a handoff from a predecessor task are all excess motion. Additionally, when a task is multitasked, time is required for setting the task down and/or picking it back up. This time is all non-productive from the task's viewpoint and is therefore waste ([Fig. 6-12](#)).

The sixth category of waste is *non-value-added processing*. This category can include inserting excessive or redundant reviews and sign offs. It also includes the situation where resources are required to accomplish additional tasks within the project that are not part of the

project, but that are included because the resource may be in a similar area of work ([Fig. 6-13](#)).

This happens frequently in software development projects where, in making a change in a part of the operating program for the project, the resources are asked to update the programming in the same part of the code for an additional need that is not associated with creating value for the project at hand.

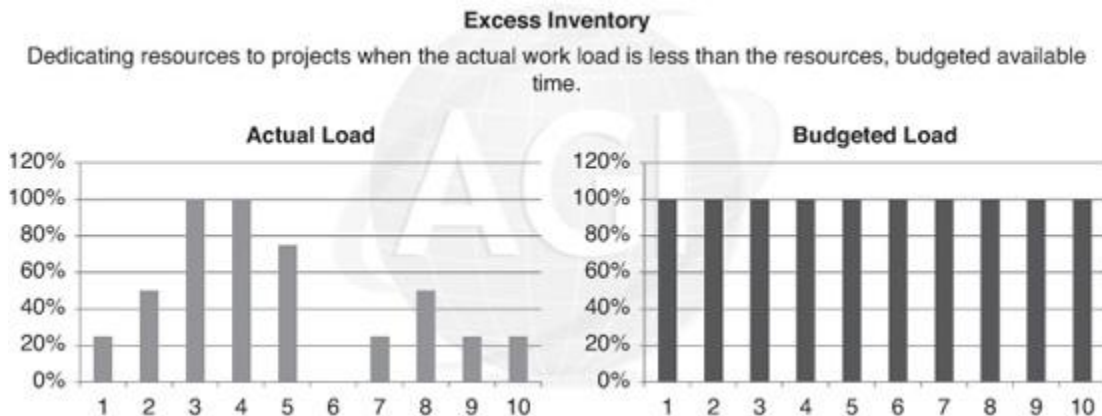


FIGURE 6-11 Excess inventory. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

Excess Motion
Unnecessary set up and set down of a task

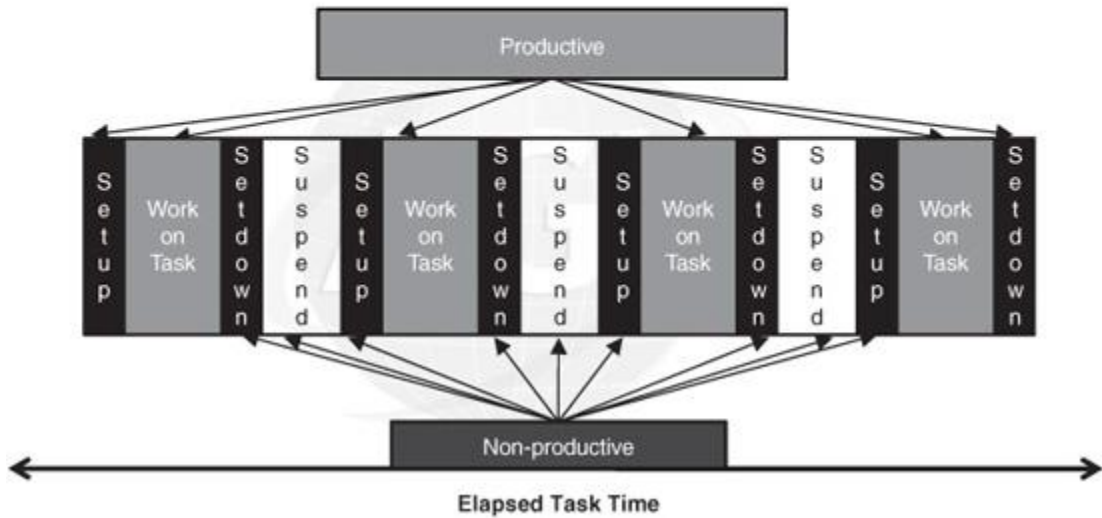


FIGURE 6-12 Excess motion: Unnecessary set up and set down of a task. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

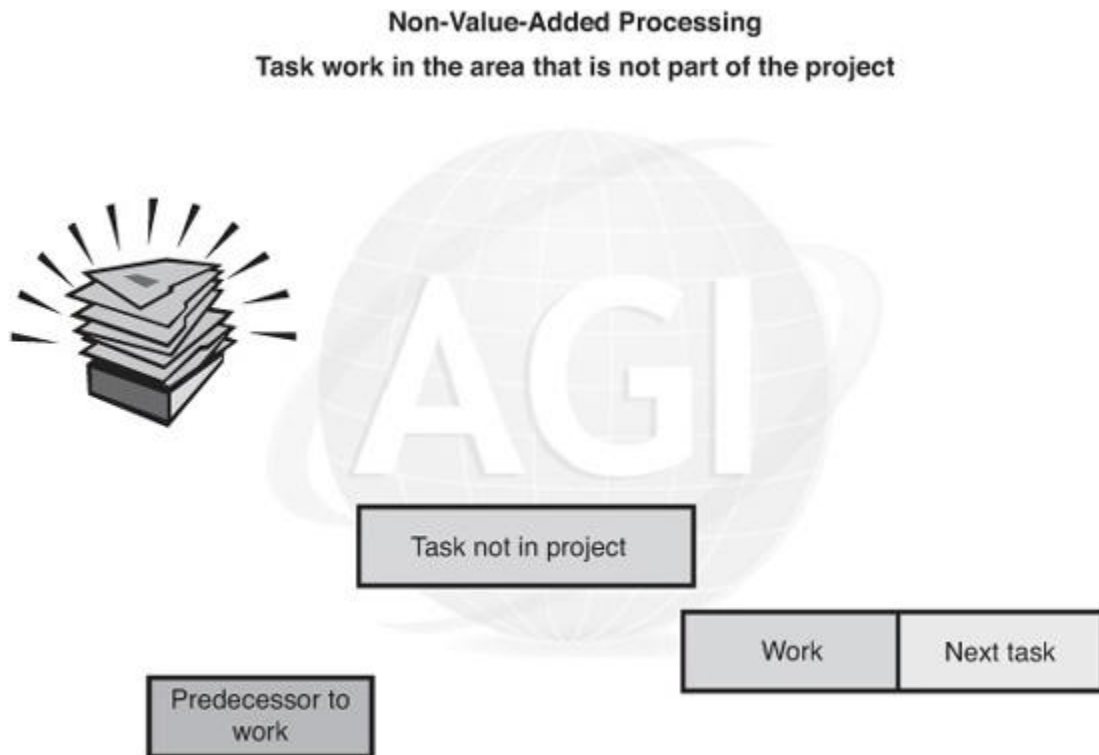


FIGURE 6-13 Non-value-added processing. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

The seventh category of waste is *defects*. Defects can take many forms, from wrong, missing, or incomplete information to handing off a task that does not meet its exit criteria. The defects category also captures the situation when variability is not addressed in the project when it first occurs. The later the variability is discovered, the more time and task areas will have to be reworked, creating waste of time ([Fig. 6-14](#)).

The eighth category of waste is *underutilized resources*. In many project environments, within the same skill set, there are “go-to” people. Everyone wants them on their tasks and in their reviews.

In [Fig. 6-15](#), the load for the two blue skilled resource is 100 percent, but when we look at the load by individual person, one is loaded 170 percent, while the other is loaded only 30 percent and is underutilized.

Pursuing Perfection

The fifth principle of Lean that Womack and Jones (1996) cite is the *pursuit of perfection*. Lean practitioners are asked to visualize the “perfect” process. No matter how much you improve a process to make it leaner, there are always ways to continue to remove waste by eliminating effort, time, space, and errors. There are six key ways to pursue perfection in projects. They are

1. Address variability at the earliest point in the project.
2. Plan how you desire to do the project (not the way you think will fit or have always done it).

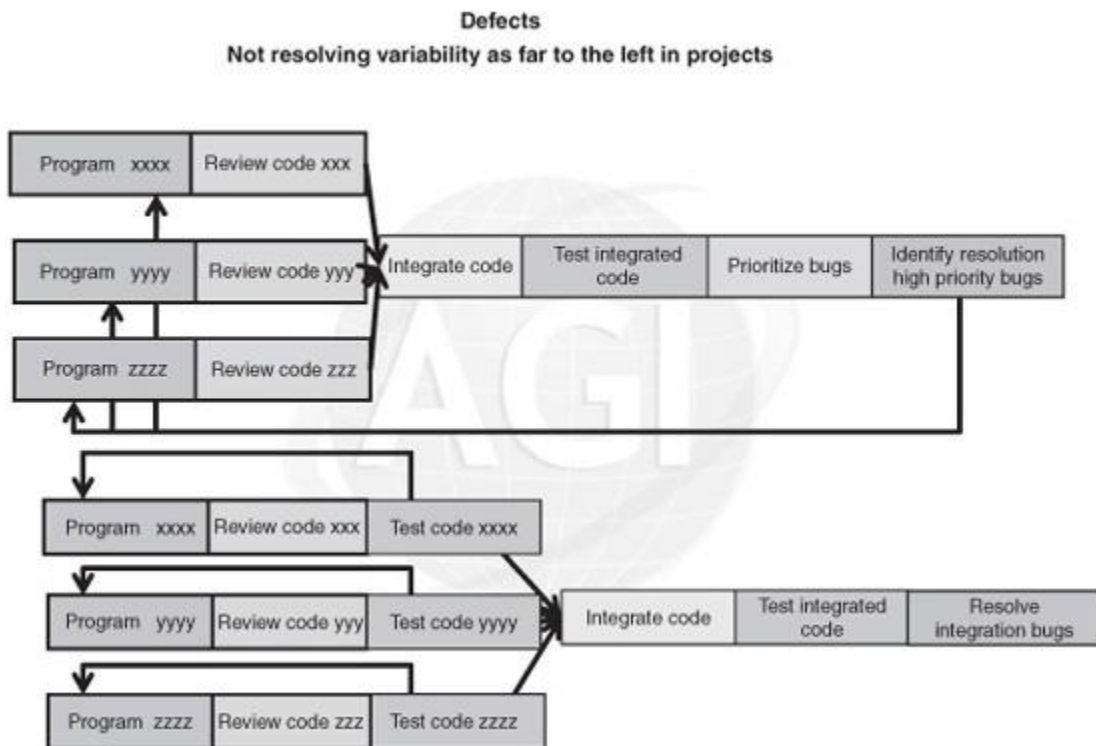


FIGURE 6-14 Not resolving variability. ©1991–2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

Underutilized People
The go-to people selected over others
with the same skills.

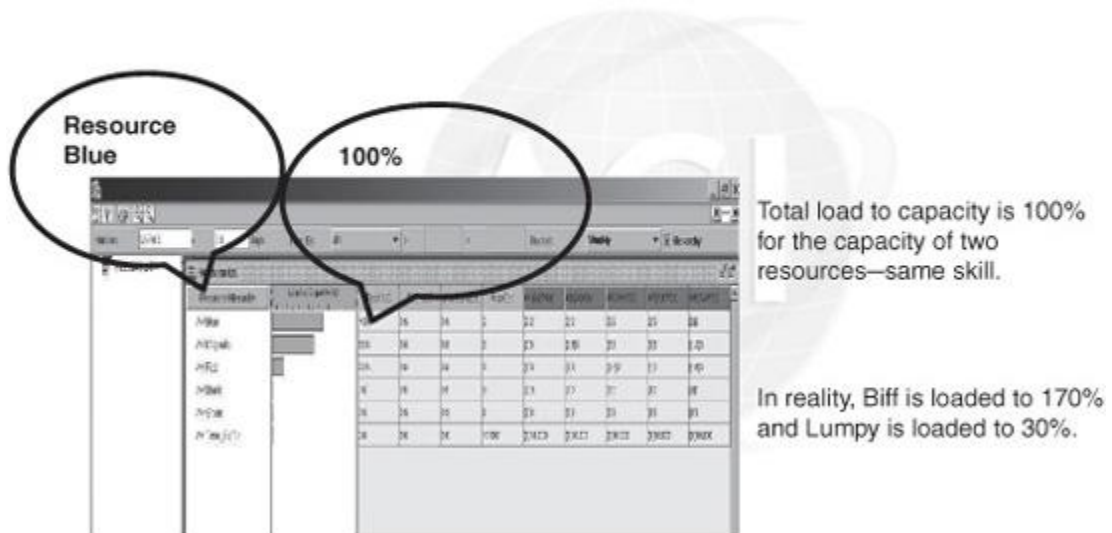


FIGURE 6-15 Underutilized people. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

3. Don't commit to a work-around until you see if one is needed (or can check for any negative consequences of the work-around).
4. Template best project practices into a PERT or network diagram and use for all like projects.
5. Apply project-based risk management to the project prior to commencing the project.
6. Monitor "actual to plan" for what causes project cycle time to expand or contract and reduce all sources of variation (in the right order).

How does one reduce variation in an environment where each project and task appears to be unique? Again, one has to understand that variability takes different forms in projects. There are four types of variation that can be

addressed in the project environment: project scope, task, iteration, and resource-to-resource.

One major cause of project scope variation (scope creep) is not gaining full consensus on the project scope upfront. This occurs either through not having all of the key stakeholders in the room ahead of time or by not pursuing the correct questions with them. By having the correct participants specify upfront their problems to be addressed by the project, the group can agree on what really needs to be delivered to solve the problems—the resulting project objectives. Additionally, with the right participation and the problem better understood, the correct scope needed is better able to be identified upfront, thus reducing scope creep.

Since tasks in projects are most often unique to the type of work of the specific project and therefore will not necessarily repeat from project to project, we often need to focus on understanding which tasks have the greatest potential for possible variability—the largest spread (longest tails).

Task variability ([Fig. 6-16](#)) refers to the difference in time between the task going pretty well (aggressive but possible) and the potential for things to go wrong (highly probable). The larger potential variations can be addressed upfront to minimize their occurrence by inserting predecessor tasks, utilizing different methods, or preventing variability from flowing to the task from an upstream predecessor task.

Iteration variability can affect the ability of a project not only to go faster, but also to be accomplished reliably. In product development, it may be referred to as a loop. “A project may go through the loop multiple iterations—testing, retesting . . . analysis, reanalysis . . . query, requery and so on. It cycles through until we have the results the client contracted us to achieve and- or -until we know everything we need to know.” (Jacob, Bergland, and Cox, 2009, 61). Iteration variability should be identified during planning and

checked to see if it is a result of waste due to defects or transportation. If so, try to reduce the iteration. Quantify the impact of the repeatable variation within and across projects for a possible LSS event.

Many in project environments believe that there are significant differences in time taken between skilled resources within a group—*resource-to-resource variability*. This variability is often reduced when each resource is allowed to focus on a task without multitasking. If resource-to-resource variability remains, capture and address appropriate resource-to-resource variation with mentoring provided during project execution.

At the end of each project completion, the team should perform an analysis of the variability identified before execution versus the variability actually incurred. Categorizing the tasks by which task met or beat their more optimistic (aggressive, but possible) times allows the organization to better establish the times for planning and for protecting against variability in the next project. By categorizing the tasks met or exceeding the highly probable estimate of variability, analysis should be done on the impact of these more variable tasks that require project recovery actions. Which type of variation is hurting the project the most? From which tasks or resource types? Analyze those items that provide opportunity for system-wide lead-time reduction by addressing the variation through LSS events.

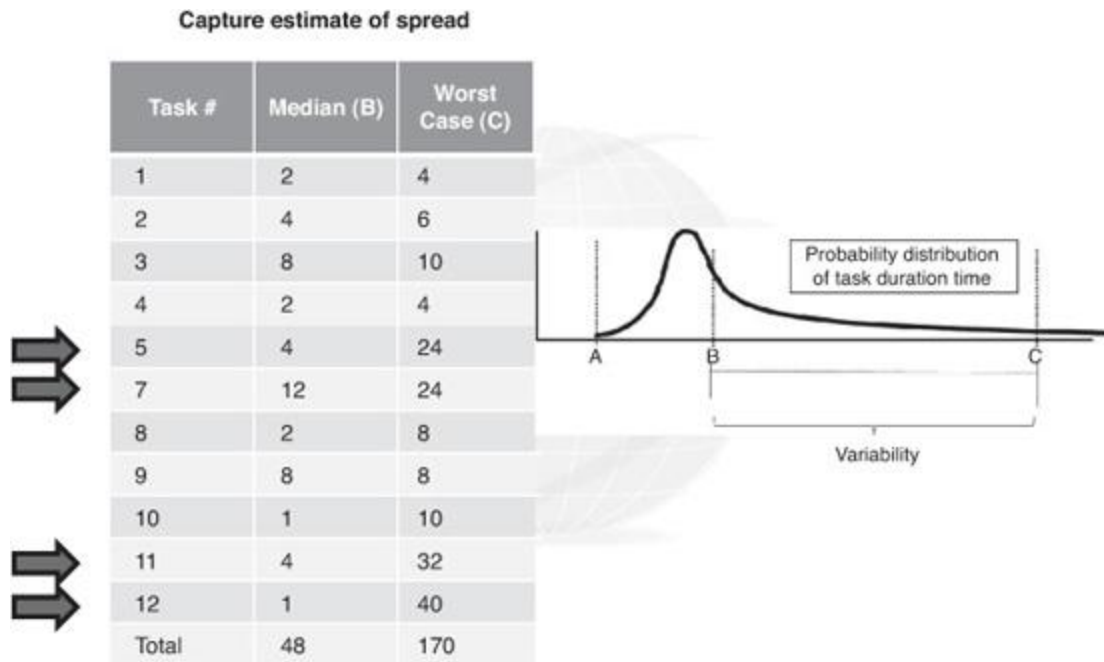


FIGURE 6-16 Task variability in a project. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

Leaning Traditional Project Management

Traditional Project Management will need some refinements to become Lean—allowing more projects to reduce their cycle time. There are improvements already developed through the TOC Project Management (TOC PM) methodologies. Through TOC PM, the alignment of the system of systems is already established. A portfolio's work is pipelined (synchronized) in accordance with capacity of the organization. More realistic, but shorter schedules are created by first planning the work as more of the perfect process called Network Building. Inherent in this process is the identification of variability. The assignment and execution of tasks based on the synchronized project's Critical Chain schedules allows the work of the project to flow value towards the customers ([Fig. 6-17](#)). Capturing

actual to plan and focusing improvement allows more effective utilization of resources to the tasks that drive project cycle time. Through over 20 years of applying the techniques of TOC PM to different project environments, we see the value of driving out waste—faster and faster projects, less compromises, and more capacity freed up.

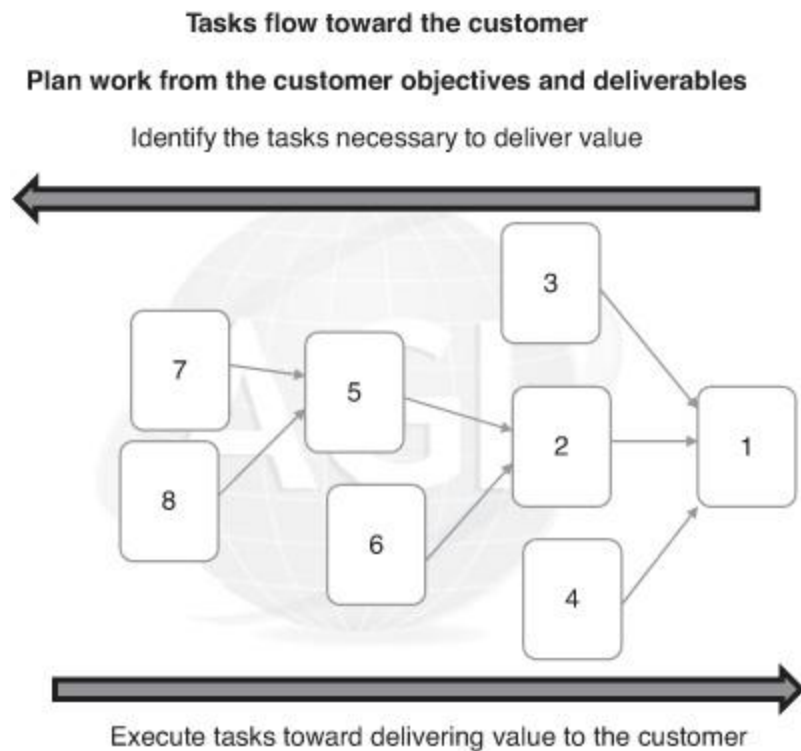


FIGURE 6-17 Tasks flow toward the customer. ©1991-2010 Avraham Y. Goldratt Institute, LP. All rights reserved.

References

Jacob, D., Bergland, S., and Cox, J. 2009. *VELOCITY: Combining Lean, Six Sigma and the Theory of Constraints to Achieve Breakthrough Performance*. New York: Free Press.

Ohno, T. 1988. *Toyota Production System: Beyond Large-Scale Production*. New York: Productivity Press.

Saxe, J. G. 1873. "The Blind Men and the Elephant" in *The Poems of John Godfrey Saxe*. Complete Edition. Boston, MA: James R. Osgood and Company, 77-78.

Womack, J. P. and Jones, D. T. 1996. *Lean Thinking*. New York: Free Press.

About the Author

Since 1986, AGI-Goldratt Institute has enabled organizations to better align the way they operate with what they are trying to achieve—strategic bottom-line results.

AGI is the birthplace of constraint-based techniques and solutions for business success. Many organizations and consultants trace their roots back to AGI not only for TOC, but also for how TOC integrates with other improvement methods.

AGI provides its clients with rapid, bottom-line results with what it calls *VELOCITY*—a powerful business approach combining speed with direction. **VELOCITY** consists of three pillars: TOC, the system architecture; TOCLSS, the focused improvement process; and SDAIS, the deployment framework.

SDAIS (Strategy-Design-Activate-Improve-Sustain) begins with creating and then executing the strategic roadmap to ensure that business processes are designed and aligned to achieve the **strategy**. Once **designed**, the business processes are **activated** to allow the organization to operate in a stable, predictable manner with less investment and organizational churn.

Once stable, focused system **improvements** are applied to increase sustainable bottom-line results. Execution

management tools and transfer of knowledge enable each aspect of SDAIS and serve as the foundation for self-sufficiency and **sustainment**.

AGI has expertise in TOC, TOCLSS, and SDAIS, with years of experience adapting each of these elements to meet the unique needs of its clients, regardless of size or industry.

AGI excels at leading organizations through successful business transformations by providing business assessment, implementation support, execution management tools, training, and mentoring.

We are motivated by making the complex manageable and enabling our clients' self-sustaining success.

SECTION III

Drum-Buffer-Rope, Buffer Management and Distribution

[CHAPTER 7](#)

[A Review of Literature on Drum-Buffer-Rope, Buffer Management and Distribution](#)

[CHAPTER 8](#)

[DBR, Buffer Management and VATI Flow Classification](#)

[CHAPTER 9](#)

[From DBR to Simplified-DBR for Make-To-Order](#)

[CHAPTER 10](#)

[Managing Make-To-Stock and the Concept of Make-To-Availability](#)

[CHAPTER 11](#)

[Supply Chain Management](#)

[CHAPTER 12](#)

[Integrated Supply Chain](#)

This section makes clear the fact that Constraints determine the performance of a system. Identifying the constraint, then knowing how to manage it and the activities it depends on to maximize constraint performance and thus organization performance is explained. With Drum-Buffer-Rope (DBR), we see how to pace work flows through a system by timing release of new work and buffering system

leverage points for statistical fluctuations, all in a way to maximize Throughput and minimize flow time. The newer concept of Simplified Drum-Buffer-Rope is explained in detail as is the framework for a pull material requirements planning (MRP) system based on Theory of Constraints (TOC) and lean concepts. We see how Buffer Management provides clear focus on priorities for expediting to prevent production delays and spotlights the best applications for improvement measures, pointing to specific areas where improvement will do the most good and count toward the bottom line. It thus becomes a centerpiece for implementing a process of ongoing improvement. Buffer Management works in job shops, assembly plants, supply chains, projects, paperwork flows, and many other environments.

CHAPTER 7

A Review of Literature on Drum-Buffer-Rope, Buffer Management and Distribution

John H. Blackstone Jr.

Introduction

This chapter is the lead chapter in a section on the Theory of Constraints (TOC) approach to production and inventory planning and control. The focus is to highlight literature on Drum-Buffer-Rope (DBR) scheduling and execution and control of that schedule through Buffer Management. Today, TOC experts believe that Buffer Management is a necessary condition for an effective Drum-Buffer-Rope system. This scheduling and control mechanism has been extended across supply chains to pull inventory to consumers. This extension of TOC into supply chains is known as Rapid Replenishment.

This chapter reviews articles describing the nature and application of DBR, Buffer Management, and the TOC approach to replenishment.

In the *TOCICO Dictionary* (Sullivan et al., 2007, 18), *drum-buffer-rope* is defined as “(t)he TOC method for scheduling and managing operations. Usage: DBR uses the following: (1) The drum, generally the constraint or CCR, which processes work in a specific sequence based on the customer requested due date and the finite capacity of the

resource; (2) Time buffers which protect the shipping schedule from variability; and (3) A rope mechanism to choke the release of raw materials to match consumption at the constraint." (© TOCICO 2007, used by permission, all rights reserved.)

¹ The *TOCICO Dictionary* (Sullivan et al., 2007, 7) defines "*capacity constrained resource (CCR)—Any resource that, if its capacity is not carefully managed, is likely to compromise the throughput of the organization*" (© TOCICO 2007, used by permission, all rights reserved.).

Copyright © 2010 by John H. Blackstone Jr.

The concepts underlying DBR were first laid out by Goldratt² (1984) in *The Goal*, although the actual terminology first appeared in Goldratt and Fox (1986), *The Race*.

DBR is the scheduling and control mechanism used to implement Theory of Constraints in a service or production facility. The term comes from the concept that the slowest station in a facility (or the market if all workstations have extra capacity) must set the pace for all the other stations, or else inventory will grow unchecked at the slower stations. This *slowest station (or the market) that sets the pace for the shop* is called the *drum*. The *buffer is material (represented as time) upstream of the drum making sure that it is never starved for work*. The *rope is a signaling mechanism from a buffer to the gateway station pulling material into the shop at the rate the drum completes material*.

The purposes of this chapter are sixfold. First, the precursors to TOC scheduling are described. Second, a review and critique of the literature on DBR scheduling is presented. Third, special cases such as free goods (when the market is the constraint), re-entrant flows, and remanufacturing are discussed. Fourth, a review and

critique of the literature on Buffer Management are presented. Fifth, the literature on TOC replenishment is reviewed and critiqued. Sixth, some suggested problems within DBR as presented in the academic literature are discussed. After the introductory section, the chapter is organized to follow the outline of the purposes and conclude with a summary and recommendations for future research. Two overarching objectives of this chapter are to provide academics a suggested framework and to provide information so that they and others can build a solid foundation of principles for further simulation and case study research.

Literature on Precursors of TOC and DBR

TOC represents one in a long line of improvements to manufacturing operations, which include interchangeable parts, the moving assembly line, and assembly line balancing.

Historical Developments Preceding TOC

There were a number of key developments preceding TOC. Without attempting a review of the development of the Industrial Revolution and the Information Age, I present here some highlights including the development of interchangeable parts, the creation of the moving assembly line, assembly line balancing, *Just-in-time* (JIT) planning and control systems, and the *Optimized Production Technology* (OPT®)³

Interchangeable Parts

Eli Whitney, the cotton gin inventor, is usually credited with developing interchangeable parts for his contract to make muskets for the U.S. government in the late 18th century. However, a large number of companies contributed to the development of interchangeable parts. Those who credit Whitney with the innovation note that, as a firm doing business with the U.S. government, Whitney's firm was required to make his innovation available to the armories at Springfield and Harper's Ferry, Virginia. Both of these armories made substantial use of interchangeable parts.

Conti and Warner (1997) quote Boorstin (1965) as describing interchangeable parts as "the greatest skill saving innovation in human history," enabling workers without specialized skills to make complex products. Conti and Warner date the history of interchangeable parts back to the mid-16th century, when the Venetian Arsenal used standardized parts in shipbuilding.

Advantages of Interchangeable Parts Interchangeable parts drove down unit costs and made available a large stock of replacement parts so that a failed unit could easily be repaired.

Disadvantages of Interchangeable Parts Initially, the items made from interchangeable parts lacked variety and thus failed to meet market demand. These finished goods also lacked the flair and uniqueness of a piece made by an artisan. Introducing new products was problematic because of the difficulty of making all new machine tools.

The Moving Assembly Line

To achieve a high-volume mechanical assembly line requires reliable precision equipment and standardized shop practices (Heizer, 1998). In August 1908, while still producing the Model N, Henry Ford hired Walter Flanders

who brought to Ford a much needed knowledge of machinery, layout, and production methods (Sorensen, 1956). The initial moving final assembly line proved so successful that three of them were built in the fall of 1913 (Heizer, 1998).

Advantages of the Moving Assembly Line By moving the work to the worker, the worker did not have to move all of his tools and materials to the work. This saved a great deal of time and made the assembly process much cheaper.

Disadvantages of the Moving Assembly Line Because the assembly line moved at a specific pace, the automobile chassis was in a given station for only a certain number of seconds. If any problem arose, that particular chassis could not have the operation completed before the chassis moved out of the area. This problem necessitated a “fix-it” station at the end of the assembly line, where automobiles with problems that occurred during assembly were completed.

Assembly Line Balancing

In designing an assembly line, the number of workers, and hence the direct labor cost, is minimized if every worker or station has an equal amount of work. If every station has an equal amount of work, the number of stations is minimized. Thus, a common field of study regarding assembly lines has long been assembly line balancing. Amen (2000) developed a list of heuristics for assembly line balancing. He later (2001) performed a study of the comparative performance of these methods. Becker and Scholl (2006) extend the discussion to include U-shaped lines, as are common in JIT facilities, and mixed model lines. U-shaped lines are used primarily to produce components for JIT operations, with material entering at one end of the U and exiting at the other. Workers usually perform multiple tasks with tasks

often on each side of the U. The number of workers in the line varies by season to maintain a daily output consistent with daily sales.

Just-in-Time

The Toyota Production System (TPS) and Kanban System (Sugimori et al., 1977) were “developed by the Vice-President of Toyota Motor Company, Mr. Taiichi Ohno and it was under his guidance that these unique production systems have become deeply rooted in Toyota Motor Company...” Just-in-time is the successor of the TPS. The purpose of using JIT is to eliminate waste from processes (Hall, 1997). The name JIT is misleading because it suggests that the concept primarily involves materials arriving just in time for use. The major benefit of JIT techniques is the simplification of the processes themselves.

JIT implements a pull system of control often using cards or Kanbans to implement the pull system in which materials are replenished at approximately the same rate they are used.

The objective of JIT is to streamline a process—to change and improve the process itself, not to install a control pull system on a process undeveloped for it. Improvement is multidimensional: delivery (lead time and due date performance), cost, quality, customer satisfaction, and so on.

OPT[®]—The Precursor to DBR

DBR gradually evolved out of Goldratt’s experience with a shop floor scheduling software called OPT[®]. In his article “Computerized Shop Floor Scheduling,” Goldratt (1988) explains in detail how OPT[®] evolved. The first version of the software was basically automated Kanban. Goldratt states

that early versions of OPT[®] were such that straightforward usage was restricted to repetitive environments.

Goldratt came to realize that not all machines need to be utilized 100 percent of the time—only constraints need this. OPT[®] was reformulated to limit non-constraints to only the work necessary to keep constraints properly fed. This led to difficulty convincing supervisors of non-constraint resources to follow the schedules when these schedules called for less than 100 percent utilization. Goldratt realized that only the bottlenecks should be scheduled—other stations have excess capacity and can keep pace—and thus data accuracy was really needed only at the constraint.

The Nine OPT[®] Rules We will now list the nine OPT[®] rules (Goldratt and Fox, 1986, 179)⁴ and discuss them as special cases of mathematical programming and other methods:

1. *Balance flow not capacity.*
2. *The level of utilization of a non-bottleneck is not determined by its own potential but by some other constraint in the system.*
3. *Utilization and activation of a resource are not synonymous [sic].*
4. *An hour lost at a bottleneck is an hour lost for the total system.*
5. *An hour saved at a non-bottleneck is just a mirage.*
6. *Bottlenecks govern both throughput and inventories.*
7. *The transfer batch may not and many times should not be equal to the process batch.*
8. *Process batches should be variable not fixed.*

9. *Schedules should be established by looking at all of the constraints simultaneously. Lead times are the result of a schedule and cannot be predetermined.*

It is often counter-productive to attempt to balance capacity in order to get a flow-balanced plant. Because constraints determine system performance, constraints should have a buffer of material (represented as a *time buffer*) upstream of them to protect them from out-ages occurring upstream. This buffer will disappear as it is used to protect from outages. If the upstream workstations have the same capacity as the constraint, the buffer can never be rebuilt and the constraint utilization becomes a function of the vagaries of outages of upstream stations. To balance flow, the capacity upstream of the constraint needs to be bigger than constraint capacity to rebuild the material buffer. Likewise, when stations downstream of the constraint experience outages, the constraint will eventually run out of *a place to store its output* (the *space buffer*). Stations downstream of the constraint need more capacity than the constraint to empty the space buffer as needed.

In a simple line, it is easy to see that constraints determine non-bottleneck performance. If there are two or more bottlenecks in a line, the *constraint* will be *the station with the least capacity*. Stations downstream of the constraint can process no faster than the constraint because material must pass through the constraint to get to them. Stations upstream of the constraint could work faster than the constraint, but this will build inventory at the constraint and eventually the futility of having upstream non-constraints working faster than the constraint will be recognized and the practice will be stopped.

To *activate* (a non-constraint producing more work that the constraint can process) a resource when the resulting output cannot get through the constraint is the meaning of Rule 3. Activating a non-bottleneck resource to produce

more than can be processed by the constraint does not add any value to the company.

A *bottleneck* is a bottleneck only if *it cannot keep up with market demand working 24/7*. Thus, there is no reservoir of time from which an hour lost at the constraint can be replaced. It is simply lost to the system.

It has long been known that the slowest station in a line determines output. OPT[®] extends this principle to job-shop type flows. In a job shop, the constraint may shift around somewhat as the mix of orders varies from season to season but there is generally one machine that is the heart of the plant and the reason most of the orders are obtained. This machine or work center tends to be needed on almost every job and becomes a long-term constraint on the system. Thus, even beyond simple lines, the constraint determines the output of the system.

By having the *transfer batch* (the number transferred between two stations) be less than the *process batch* (the number processed between setups), it is possible to have several stations working on an order simultaneously. This gets the order through the facility very quickly. It could be done to expedite the order. Alternatively, it could be done simply to use a short lead time as a competitive weapon in the marketplace.

Process batches should be variable, not fixed. If a product is seasonal and a shop always makes one week's worth of demand as a process batch, then the process batch will vary naturally over the course of the year. This approach allows little inventory to accumulate. If a fixed process batch large enough to cover a week's demand during the peak season was to be used, inventory covering several weeks demand would be created during the off-peak periods. Having a variable process batch makes more sense. Traditional plants may use the *Economic Batch Quantity* (EBQ) (*the number of units processed at a time to minimize setup and carrying*

costs) formula to determine a fixed process batch size, but the EBQ formula assumes a fixed demand so its use really is not appropriate in this situation.

Derivation of DBR Using the Five Focusing Steps

TOC says that constraints (anything that limits a system from achieving a higher performance versus its goal) determine the performance of a system and TOC provides methods for efficiently and effectively utilizing these constraints. Since it is not the main topic of this chapter, here I will present only a key definition and the *Five Focusing Steps* (5FS) without elaboration or delving into ramifications. There is expanded coverage of 5FS in [Chapter 8](#) and elsewhere in the book.

The 5FS are as follows:

1. *Identify* the system constraints.
2. Decide how to *exploit* the system constraints.
3. *Subordinate* everything else to the above decision.
4. *Elevate* the system constraints.
5. If, in the previous steps, the constraints have been broken, *go back to Step 1*, but don't let inertia become the system constraint (Goldratt, 1988).

In Step 1, a company defines its drum. In Step 2, it develops buffers at shipping and the internal resource constraint if it exists. At Step 3, the rope is tied between the buffer and material release to maintain the constant buffer.

A number of articles have discussed 5FS. These include Mabin and Davies (1999), Ronen and Spector (1992), Jackson and Low (1993), Politou and Georgiadis (undated),

Mabin and Davies (2003), and Trietsch (2005). In addition, Gupta et al. (2002) introduced a series of simulation models that were run with each successive model introducing another step.

Jackson and Low (1993) note that an important contribution of constraints management is the focus it provides the entire organization. When everyone understands the vital role the constraint plays in the organization, everyone measures their actions according to the effect on the constraint and thus the total productivity of the system.

Scheduling the Resource Constraint

In TOC, all workstations work to maintain the schedule set at the constraint resource. Goldratt (1990) describes how this schedule is derived in *The Haystack Syndrome*. For each order, we have the due date of the order. We also have an estimate of the time it will take for the order to move from the constraint resource to the shipping dock—the shipping buffer. Scheduling the resource constraint involves loading each job onto the constraint, a shipping buffer time before its due date, and resolving any timing conflicts. The Avraham Y. Goldratt Institute produced a set of production simulators (a Windows version is provided in Goldratt, 2003b) to teach potential users constraint-scheduling concepts.

The article by Schragenheim and Ronen (1990) is the most often cited description of how DBR scheduling works. They list three steps: (1) schedule the constraint, (2) determine the buffer sizes, and (3) derive the materials release schedule according to steps (1) and (2). Schragenheim and Dettmer (2001) and Schragenheim, Dettmer, and Patterson (2009) provide perhaps the most in-depth discussion of DBR, including a special case called

simplified DBR, and such issues as multiple constraints, moving bottlenecks, multiple operations occurring at the bottleneck, and other complications. *Simplified DBR (S-DBR)* assumes that the market is the constraint and therefore uses only one buffer—the shipping buffer (frequently called the *production time buffer*). Of course, if there is an internal constraint, material will naturally accumulate upstream of the constraint establishing a de facto constraint buffer.

Scheduling Non-Constraints

The pure DBR methodology does not develop a formal schedule for non-constraints. Rather, the rope determines when material is to be released to the first station on a routing and material is allowed to flow naturally between workstations. If decisions made by workstation supervisors result in a hole deep in the buffer, then expediting by using small transfer batches to achieve overlapped operations at a few stations may be needed to get material into the buffer in time to avoid the hole reaching the buffer origin (starving the constraint).

The individual work center (non-constraint) supervisor is advised that when a hole deep in the buffer appears, he or she should schedule the missing job first. If there are no significant holes in the buffer, he or she is free to run any job next. The supervisor might choose a job because of a short, sequence-dependent setup time, for example. Many academics are uncomfortable with this informal, ad hoc, logic for dispatching at non-constraints. Some researchers have developed alternative mechanisms for scheduling non-constraints.

Protective Capacity

TOC breaks capacity at non-constraints into three categories⁵: (1) productive capacity, (2) protective capacity, and (3) excess capacity. Productive capacity is that capacity equal to the constraint's capacity—the ability to produce the number of units that the constraint can produce. Protective capacity is capacity needed to restore buffers to their ideal state after a disruption—to refill the time buffer that has become depleted or to empty the space buffer that now has material awaiting processing downstream of the constraint. This restoration of the ideal state of the buffer needs to be done quickly, before another disruption occurs. Excess capacity is capacity over and above productive and protective capacity.

Protective capacity is one of the most vital aspects of DBR because if there is insufficient protective capacity, then the buffer cannot be refilled quickly enough when the buffer is low and thus the drum is vulnerable to possible *starvation*⁶ by upstream stations or *blocking* by downstream stations. Since an hour lost at the drum is an hour of lost output if the drum is a resource constraint, downtime at the drum can be extremely expensive. Protective capacity is idle when the buffer is in an ideal state and needs no restoration. The non-constraint station uses only enough capacity to produce at the drum's pace. However, once a buffer leaves its ideal state, all affected non-constraints must use their protective capacity to restore the buffer to an ideal state before some other problem threatens to idle the drum.

Of course, in a deterministic environment there would be no need for protective capacity because a constant amount of inventory would be held in the time buffer. An issue related to the establishment of a DBR system is "How much protective capacity is needed and how should it be arranged?" There have been only a few studies of this issue. This issue is especially important if there are *capacity constrained resources* (CCRs) in the system. Recall, a CCR is

defined by the *TOCICO Dictionary* (Sullivan et al., 2007, 7) as “any resource that, if its capacity is not carefully managed, is likely to compromise the throughput of the organization.” (© TOCICO 2007, used by permission, all rights reserved.)

Literature on DBR Scheduling

In discussing literature on DBR, I first present some overview articles that principally discuss the 5FS or the nine OPT[®] rules. Then I move to simulation models and case studies divided by VAT classification. V, A, and T represent types of plants with *V-plants* dominated by divergences in flows, *A-plants* dominated by assembly operations, and *T-plants* experiencing a huge increase in variety in the final operations. After the sections on V-, A-, and T-plants, I present those simulations and cases that could not be assigned a specific VAT class.

Overviews

When TOC first appeared, total quality management and JIT were also gaining popularity. Because Goldratt was doing most of his information transfer via workshops and his books *The Goal* and *The Race*, many people lacked a true sense of what TOC entailed. A number of people sought to fill this relative void by introducing articles covering the 5FS or the nine OPT[®] rules, and especially DBR and Buffer Management. Some also reported on DBR implementations. Because the articles are broader than case studies of a single implementation, this section of the chapter was developed to gather these broad overviews.

Cox and Spencer (1998) devote a chapter of *The Constraints Management Handbook* to the DBR scheduling

method. Throughout the chapter, they give a detailed three-product, five-work-center example, showing how to develop a schedule for the drum and for shipping. They also present a section on Buffer Management and a section on how DBR works within a *material requirements planning* (MRP) system. Overall, this is an excellent short summary of DBR.

Mabin and Balderstone (2000) present a book containing over 300 abstracts for books and papers on TOC published prior to 2000. They present a tree showing all aspects of TOC, with DBR belonging to a branch on production management. Only a few of the abstracts relate to DBR, including Atwater and Chakravorty (1994) who present a simulation study of the importance of protective capacity. Betz (1996) presents a study of an implementation at Lucent Technology. Coman et al. (1996) discuss the successful implementation of DBR at an Israeli electronics firm. Conway (1997) presents a concern over DBR scheduling the constraint carefully and non-constraints loosely as previously described by Simons and Simpson (1997).

Danos (1996) discusses how an implementation of DBR software increased profits by 300 percent at one company. Demmy and Demmy (1994) present a novel use of DBR by a photographer (treating himself as the constraint) in scheduling students to have their pictures taken for their yearbook. Demmy and Petrini (1992) describe the successful implementation of DBR to control aircraft maintenance within the Air Force Material Command. Duclos and Spencer (1995) use a simulation model of three different environments to show how DBR produces significantly better results than MRP in a hypothetical company. Fawcett and Peterson (1991) include DBR in a discussion of manufacturing-related aspects of TOC. Fry (1990) discusses an important aspect of buffers—the impact of *work-in-progress* (WIP) inventory on lead times. Because most of the time that a part spends in a facility is waiting for

service rather than being serviced, there is a strong correlation between WIP and lead time. In a follow-up article, Fry et al. (1991) discuss the implementation of DBR to control lead time. Gardiner et al. (1992) provide a comprehensive overview of DBR and Buffer Management. Gardiner et al. (1994) present a brief discussion of DBR and Buffer Management in discussing the evolution of TOC.

Grosfeld-Nir and Ronen (1992) discuss the application of OPT[®] to the single-bottleneck problem. Lambrecht and Alain (1990) present the results of a simulation comparison of JIT and DBR. In an earlier paper, Lambrecht and Decaluwe (1988) show that DBR is more robust than JIT in managing bottlenecks. Pinedo (1997) provides a second commentary on Simons and Simpson (1997) praising the overall article but raising an issue of lack of comparison with other software. Radovilsky (1994) uses queuing theory to estimate the size of time buffers in DBR (Goldratt and others suggest using an amount equivalent to a portion of the existing lead time). Radovilsky (1998) presents a follow-up to the initial article, also estimating initial time buffer size using queuing theory. It should be noted that Buffer Management would be used to adjust this initial estimate based on whether too much or too little material is present in the time buffer.

Reimer (1991) outlines DBR and discusses it with a modified MRP framework. Schragenheim and Ronen (1990; 1991) are discussed; these articles were discussed at length earlier in the chapter. Russell and Fry (1997) discuss order review/release mechanisms that could be used to fill the function of the rope and discuss lot splitting into several transfer batches as an expediting methodology. Schragenheim et al. (1994) discuss modifications of DBR for use in process industries. Simons and Simpson (1997) present a concise history of the evolution of DBR and the algorithm in detail, and relate the algorithm to alternative methods.

Spearman (1997) gives both positive and negative comments on TOC and the Goal System software. Spencer (1991) discusses the basic theory behind DBR and how to marry DBR and MRP II. Spencer and Cox (1994) discuss the distinctions between OPT[®] and TOC. Spencer and Wathen (1994) present a case study of service functions at Stanley Furniture including an implementation of DBR. Stein (1996) includes a discussion of the advantages of DBR and dynamic buffering in a generalized manufacturing situation. Umble and Srikanth (1995) include a thorough discussion of DBR in their pioneering book, *Synchronous Manufacturing*.

Wolffarth (1998) presents practical lessons learned from an implementation of DBR within an Enterprise Resources Planning (ERP) system. Yenradee (1994) presents a case study from a battery factory using a manual DBR system in conjunction with the nine OPT[®] rules. Mabin and Balderstone (2000) also present a list of 34 books that had been published on TOC by 2000.

There are several overviews of TOC including some discussion of DBR published since the Mabin and Balderstone (2000) book appeared. Rahman (1998, 337) states that TOC contains two major components—the logistics paradigm, including DBR, and the Thinking Processes, which he calls a “generic approach for investigating, analyzing, and solving complex problems.” He includes the 5FS, the nine OPT[®] rules, and the definitions of the three operational measures (Throughput, Inventory and Operating Expense). He also includes a table of 139 articles and conference proceedings broken down by year and journal. Gupta (2003) provides an overview that relies heavily on both Rahman and Mabin and Balderstone as an introduction to a special issue of *International Journal of Production Research*. Watson et al. (2007) update the comprehensive discussion of the evolution of TOC previously discussed in Gardiner et al. (1994).

Boyd and Gupta (2004) give an excellent overview of TOC, comparing its philosophy to several somewhat similar philosophies but giving only a rudimentary overview of DBR.

Applying DBR to Different Types of Facilities: VATI Analysis⁷

The *TOCICO Dictionary* (Sullivan et al., 2007, 51) defines “*VATI analysis—The stratification of operations environments into four generic types referred to as: V, A, T, and I. Each environment has an inherent set of undesirable effects that, properly understood, make operations management easier. Each type is named for the letter that resembles a diagram of the logical flow (not the physical flow) of materials. Usage: A single plant may be a combination of more than one type.*” (© TOCICO 2007, used by permission, all rights reserved.)

Umble and Umble (1999) discuss VAT analysis; that is, classifying plants as one of these three types and recognizing that certain characteristics are common to each type. They state that VAT classification was developed around 1980 by Goldratt.

Product flow diagrams for *V-plants* are characterized by divergence points (hence the V-shape). Three characteristics are typical of V-plants:

1. The number of end items is large compared to the number of raw materials.
2. All end items sold by the plant are processed in essentially the same way.
3. The equipment is generally capital-intensive, highly specialized, and typically requires lengthy setups.

A-plants are characterized by convergent assembly points throughout the process. In such plants, a large number of purchased or fabricated component parts and materials, generally produced in a job shop environment, are combined to form subassemblies that are used to build unique end products.

T-plants are dominated by a major divergent assembly point at final assembly, where many different end items are assembled from a relatively limited number of component parts.

Umble and Umble (1999) go on to discuss the specific placement of buffers in each type of plant.

I-Plant Research

Many simulation models used to study aspects of DBR are simple I-plants. This is because issues such as what constitutes adequate protective capacity or the buffer's impact on lead time can be studied in an I-environment without complicating factors that occur in A-, V-, or T-plants.

Fry et al. (1991) simulate an I-plant to show how having little WIP at non-constraints in DBR gives strong control of lead time.

Finch and Luebbe (1995) simulate a five-station system in which the constraint moves over time because of different learning curve rates at the five stations. Because of shifting work center times during much of the simulation, there is little or no protective capacity at non-constraints. The authors conclude that there are significant interactions between learning curve effects and constraint production and that there is need for further study of this issue.

Atwater and Chakravorty (1996) simulated simple 5- and 6-station serial lines (I-lines) with disruptions created by machine breakages using balanced, JIT, and TOC configurations. They found that TOC-based lines are less

affected by variability than either balanced or JIT lines. Chakravorty and Atwater (1996) study line design for I-structures.

Kadipasaoglu et al. (2000) simulating an I-facility found that (1) when protective capacity increased from 0 to 12.5 percent, flow time decreased by about 40 percent; (2) there is a benefit to WIP level for having the constraint be the first station;⁸ and (3) non-constraint downtime and protective capacity tend to have opposite effects on flow—increasing non-constraint downtime decreases flow, which can be offset to an extent by increasing protective capacity. Betterton and Cox (2009) later studied this simulation and found that the methodology employed was not a correct implementation of DBR. First, Kadipasaoglu et al. (2000) had random arrivals released into the plant rather than using a rope to release material at the drum's pace. Second, using station 1 as the constraint, Kadipasaoglu et al. used infinite buffers at all downstream stations. Blocking can never occur with infinite buffers, so the constraint would never undergo blocking. Simulating the environment as a true DBR environment, Betterton and Cox (2009) found that some of Kadipasaoglu et al.'s findings were not correct.

Blackstone and Cox (2002, 419), using a simulated I-facility, define "*protective capacity*" as "the capacity needed at non-constraint workstations to restore WIP inventory to the location adjacent to and upstream of the constraint workstation to (create a time buffer to) support full utilization of the constraint workstation." It should be noted that the ability of downstream stations to empty the space buffer when it contains work is also protective capacity—protecting against blocking. Blackstone and Cox also show that the size of the time buffer required to adequately protect the drum is inversely related to protective capacity, a point that had been made previously by Atwater (1991).

Kim et al. (2003a) simulated a variety of flow control mechanisms within an I-line and found that, compared to output flow control and dynamic flow control, bottleneck flow control achieved greater output with less WIP while maintaining smaller lateness and tardiness of orders.

Real I-Lines I found no simulation studies or case studies of real I-lines. I think this is because even when the flow is straight line, real facilities tend to have multiple products that diverge into various configurations as they travel down the line. That is, they are V-plants, not I-plants.

V-Plant Research

Simulations of Real V-Plants Vaidyanathan et al. (1998) describe the simulation of a coffee production facility having moving CCRs. The simulation model was used to develop a schedule for this V-plant. The simulation showed that output could be increased by approximately 40 percent by using the simulation model to develop the schedule.

Hasgul and Kartal (2007) used the Wagner-Whitin algorithm, a very sophisticated technique used to attempt an optimal sequence of jobs over a lengthy planning horizon, to schedule a simulated refrigerator plant. The portion of the company they were simulating corresponded to a V-plant. They reported achieving an average cycle time decrease from 12 days to 7 days when DBR was applied.

Case Studies of V-Plants Chakravorty (1996) reports a case study at Robert Bowden, Inc., a \$40 million sales supplier of residential and light commercial building products whose manufacturing facility is a V-plant. After implementation of DBR (which is described in the article), the average number of orders processed increased by 20 percent with no increase in staff, and expediting of orders was significantly reduced.

Rerick (1997) presents a study of semiconductor wafer manufacture at Harris Corporation, which reduced cycle time by approximately 50 percent while almost doubling output. Wafers were made for automotive, telecommunications, and computer markets. A control point was selected to implement a DBR system.

Huang and Sha (1998) use a hybrid DBR/Kanban system to model a wafer fabrication facility through a simulation model. Kanbans, which pull material forward station-by-station, somewhat override the purely informal DBR approach to non-constraint dispatching. Huang and Sha also attempt to determine the optimal size of Kanbans in such a system.

Hurley and Whybark (1999) studying a simulated V-plant correctly point out that variance reduction can reduce the need for protective capacity and protective inventory.

Chakravorty (2000) presents a second case study of DBR at Robert Bowden, Inc., emphasizing the fact that it is a V-plant. V-plants running DBR have not received a great deal of attention in the literature. The plant used two buffers—constraint and shipping. Between 1996 and 1999, annual sales in units increased from approximately 58,000 to over 80,000 while the number of workers only increased from 12 to 16. During the same period, the stock of finished goods was reduced from 3800 to 1325, while late orders decreased from 19 to 7 percent.

Frazier and Reyes (2000) present a detailed description of how DBR was applied to the Dallas, Texas, plant of a company manufacturing cable and telecommunication equipment in a V-plant. After three months, WIP decreased to one-third of its previous level, raw materials inventory value decreased by approximately 30 percent, and percent on-time completion of jobs increased by more than 30 percent.

Schaefers et al. (2004) report the implementation of DBR in a facility that buys large rolls of metallic sheets and cuts

them into smaller coils with less width and length. This appears to be a V-plant. The firm is a make-to-order (MTO) operation with no internal constraint so it used the shipping schedule as the drum (S-DBR). Before implementation, lead time varied from 21 to 182 days. After implementation, it was a stable 10 days. Customer service level increased from 34 to 87 percent. The exact change in profitability was not reported, but the authors did say that the facility changed from losing money to making money.

Belvedere and Grando (2005) report on a DBR implementation at an Italian chemical company producing dyes and pigments. Because the main raw materials are natural products, it was difficult to obtain the desired color precisely. The solution would be diluted and color-tested repeatedly, causing a dilution and the sample-testing department to be the constraint. In two years, the DBR led to a decrease in raw materials and finished goods inventory and to an increase in the number of stock turns, which almost doubled between 1999 and 2001.

Umble, Umble, and Murakami (2006), noting the lack of case studies from Asian implementations, report a case involving Hitachi Tool Engineering, a Japanese tool-engineering firm employing approximately 1100 people. They describe the plant they studied as a V plant. In addition to implementing DBR, the firm implemented some TOC thinking processes. A simple DBR system was set up using three shelves at the bottleneck with each shelf containing a day's work for the bottleneck. The authors report that this was adequate to buffer the bottleneck and to subordinate other resources to the bottleneck's schedule.

A-Plant Research

Simulation of Hypothetical A-Plants In this section, I discuss simulations of hypothetical lines that appear to me

to be A-plants. Unless specifically mentioned, the authors did not specify the plant type using the VATI breakdown.

Taylor (1999) simulated a traditional push (MRP) system versus a pull (JIT) and “hybrid” (DBR) system regarding their impact on financial measures. His simulation model appears to be an A-plant. It contained 29 stations. Independent variables included buffer size and location. He found that the DBR system had higher profit, return on investment (ROI), and cash flow while using considerably less inventory. The pull system placed second in financial results with the push system placing last. Taylor (2000) studied this same plant for impact on TOC operational measures such as Throughput, Inventory, and Operating Expense.

Atwater and Chakravorty (2002) found that mean flow time through a simulated system that has a jumbled flow and appears to be an A-plant decreased as protective capacity increased but at a diminishing rate as protective capacity reached 7 percent. Mean tardiness decreased in the same fashion. In their study, they varied constraint utilization from 94 to 98.5 percent. They compared releasing jobs immediately upon arrival in the system to releasing jobs according to the DBR schedule and found that while DBR had a smaller mean flow time through the system, the immediate release approach resulted in fewer tardy jobs.

Simulations of Real A-Plants Wu, Morris, and Gordon (1994) show how DBR improves makespan when compared via simulation to a traditional production control system. *Make-span* is the time from the start of processing until the final unit clears the system. The Wu et al. simulation is an A-plant based on a furniture manufacturer. They demonstrated that a Taiwanese furniture manufacturer would benefit significantly from makespan by implementing DBR. In their simulation, makespan decreased approximately 50 percent when DBR was added to the environment.

Guide (1995) presents a simulation model used to estimate ideal buffer sizes in a DBR implementation at a naval repair depot. A naval or air force repair depot completely disassembles a plane (while this may resemble a V-plant, in contrast to a V-plant parts flow down each path instead of one or the other paths), repairs or replaces components as needed (probably A-plant), and reassembles the plane (A-plant). This process is known as remanufacturing.

Steele et al. (2005) simulate a shop using both DBR and MRP. They found that DBR has much better performance and suggested the use of DBR within MRP systems. They based their simulation on a bearing manufacturer. This involved an assembly that sat atop two V-lines.

Case Studies of A-Plants Andrews and Becker (1992) present a case study of Alkco Lighting, noting “Buffer Management” as a keyword. This A-plant involves several assembly operations. Alkco changed its primary measurement from efficiency to Throughput. As a result, WIP inventory improved significantly and there was an accompanying improvement in cash flow. Prior to implementation, the company was promising delivery in 60 to 90 days, had an on-time rate of only 65 percent with 16 percent of deliveries being more than one week late. Thirty-two percent of Inventory was in finished goods. The DBR system as managed by Alkco freed up 40 percent of its total floor space. Five years into the implementation, lead time was reduced to one week, while on-time delivery increased to 98 percent, sales volume increased 20 percent, and before-tax profit increased 42 percent.

Spencer (1994) reports on improvement from Trane Co. of Macon, Georgia, where output changed from an average of three units per day to six units per day with the same workforce when DBR was implemented. At this location, Trane

assembles large air conditioners designed to cool commercial facilities.

Guide (1996; 1997) and Guide and Ghiselli (1995) present three discussions of the application of DBR in remanufacturing applications such as a military repair depot. As in Guide's (1995) simulation discussed above, this facility appears to be an A-plant (reassembly) sitting atop a disassembly operation. Disassembly is somewhat akin to a V-plant in that a single plane diverges into many components to be evaluated and repaired, replaced, or reused. However, the consensus is that a disassembly operation is different from a V-plant where a part flows to one product or another.

Luck (2004) presents an Ashridge Business School (UK) study of a supply chain centered on a manufacturing company called Remploy, which makes military garments. Remploy had two plants, a V-plant that cut material and an A-plant where sewing was accomplished. Five months into a standard DBR implementation, Throughput had increased 19 percent, output per employee was up 13.4 percent, WIP was reduced more than 50 percent, and absenteeism was down an average of 7 percent. There was some increase in transportation cost, but it was small compared to the increased profitability.

T-Plant Research

I was unable to find any simulations or studies that dealt with T-plants.

Research That Could Not Be Classified as V, A, or T

Sometimes research cannot be classified as V, A, or T. For one thing, the research may include more than one plant with the types of plant being different. Even if a single plant

is described, in many instances, there may not be enough information provided to make a reasonable conclusion on whether the plant is V, A, T, or I.

Simulations of DBR Systems

A number of individuals have simulated DBR systems, sometimes to estimate DBR's parameters, such as the time buffer, and other times to compare DBR's effectiveness with systems such as Lean or *CONWIP*. Guide (1995) experimented with different buffer sizes and Buffer Management techniques at a naval air station. Kosturiak and Gregor (1998) simulated a flexible manufacturing system (FMS) using MRP, *Load-Oriented Control* (LOC), DBR, and Kanban and found that LOC and DBR had the best performance, while DBR was easier to implement. Hasgal and Kartal (2007) combined DBR with the Wagner-Whitin lot-sizing algorithm and reduced cycle times and WIP in a simulation model compared to DBR by itself.

Kayton et al. (1997) simulated a wafer factory running DBR to better understand the impact of preventive maintenance in such a facility. They found that downtime at non-constraints can become problematic in facilities using DBR even when significant protective capacity exists.

Lea and Min (2003) simulated a seven-station, three-product line using both JIT and DBR and found that JIT had slightly higher profits and service levels. They also found that activity-based costing systems slightly outperformed traditional costing and Throughput Accounting systems.

Case Studies

Several articles present case studies of successful implementation of DBR and constraint management. These

case studies could not be classified as V, A, or T. Often the reports include multiple plants.

Gupta (1997) discusses DBR benefits to a supply chain. In 1998, Gupta discussed the need for software to implement DBR, describing some situations that are too complex for manual implementation. Koziol (1988), a manager at the Valmont plant in Brenham, Texas, discusses the successful implementation of DBR at that facility.

Spencer and Cox (1995) report a study of nine repetitive-manufacturing companies, three of which were pure JIT, three added MRP to JIT, and three added TOC (OPT[®] or DBR) to JIT. No specific improvement numbers were reported; however, they found that the existence of repetitive manufacturing does not preclude the application of any of the three production planning and control systems.

As mentioned earlier, Wolffarth (1998) presents practical lessons learned from an implementation of DBR within an ERP system. Umble and Umble (2006) describe how Buffer Management was used in two accident and emergency facilities in Oxfordshire, UK.

Guide and Ghiselli (1995) report on the implementation of DBR at Alameda Naval Air Depot. This disassembly / repair facility implemented preventive maintenance, added small transfer batches, eliminated local efficiency measures, and took other DBR-related steps. Results achieved included increasing Throughput while reducing WIP, reducing airplane turnaround times, and increasing the turns ratio. Further refinements of DBR at the facility were reported to have been planned.

Umbel et al. (2001) report a case study of DBR used within an ERP system. The case is *Oregon Freeze Dry (OFD)*, which processes products by removing water at low temperatures and pressures. A branch of OFD implemented DBR in 1997, identifying a resource constraint that was designated as the drum. ERP was implemented at about the

same time. The authors report that an ERP system makes DBR more effective. Once the drum schedule is determined, the ERP system was used to tie the rope. They state that the integration of TOC/DBR may be the key to ERP success.

Corbett and Csillag (2001) report on seven DBR implementations in Brazil. Five of the companies were multi-nationals, while two were Brazilian. Six used MRP and one used Kanban before using DBR scheduling. Average time to implement DBR was 3.6 months with the longest being 7 months. All seven companies started showing beneficial results during the implementation period. Six of the seven reported that they were satisfied with DBR. Even the one reporting dissatisfaction experienced a 50 percent drop in WIP and lead time and an increase in revenue per employee from US\$56,000 to \$64,000.

Lindsay (2005) reported on the implementation of DBR in Intel distribution centers (DCs) in an attempt to reduce order cycle time and reduce Inventory. Five DCs located in five countries have implemented DBR with an average cycle time reduction of more than 60 percent and a standard deviation reduction of more than 70 percent.

Vermaak and Ventner (undated) report the use of TOC in conjunction with computer simulation of a conveyor system in a coalmine, which resulted in an 8 percent increase in output.

Mabin and Balderstone (2003) report on an analysis of over 80 successful TOC implementations taken from a search of available literature. A portion of one of their tables reporting percentage improvements in various measures is shown below.

Measure	Number Reporting	Mean % Improvement
Lead time	34	70
Cycle time	14	65
Due date performance	13	44
Inventory	32	49
Revenue	20	83
Profitability	7	116

Huff (2001) reports that Bal Seal Engineering used DBR to increase Throughput, reduce Inventories, improve due date performance, reduce Operating Expense, and double net profit. Boeing and Rockland Manufacturing also achieved dramatic improvements relative to Throughput, Inventory, and profit.

Special Cases

The TOC literature contains a number of articles describing research that does not fall neatly into the previous categories but that are significant in their contributions to the body of knowledge. I have classified this research into the topics given in the next sections.

Free Goods

Free goods are defined as goods that do not require any resource constraint involvement in their production—they require solely non-constraints. Free goods represent an opportunity for immediate increase of Throughput with little to no increase in Operating Expense (recall that Throughput accounts for raw material expense, items that are truly variable costs). However, Chakravorty and Atwater (2005) found that DBR is very sensitive to levels of free goods. Therefore, schedules using DBR need to be aware of how orders of free goods are accepted. Specifically, they found

that the number of tardy orders increased as the level of free goods released to the shop increased. They attribute this phenomenon to the loss of protective capacity at certain non-constraint resources. Atwater, Stephens, and Chakravorty (2004) discuss the impact of free goods on system Throughput. They found three basic insights for the system they modeled. First, operating the resource constraint at a level above 98 percent resulted in erratic Throughput performance. Second, increasing protective capacity above 7 percent did not significantly improve on-time performance. That is, once a nonconstraint's capacity reached 107 percent of the constraint's capacity, further increases in capacity did not improve on-time performance. Of course, this value would be very sensitive to the number and duration of statistical fluctuations included in the model. Third, when demand for constraint goods is high, managers can improve on-time performance by limiting the orders they accept for free goods (refusing such an order would reduce future utilization of non-constraint resources).

What If the Market Is the Constraint?

What if all goods are *free goods*? That is, what if *the market is the constraint*? Pass and Ronen (2003) define a *market constraint* as a situation in which *the production capacity of every resource exceeds demand for it*; they address this issue for a high-tech firm. They note that it is usually easier to control an internal constraint that is under the roof of management than to be tossed by the ups and downs of the market. The R&D department is usually a constraint because new products are not coming out fast enough. In addition, marketing or sales may be a constraint. Since lead time is a factor of competition, small batches may be run in order to shorten lead time. This involves more setups but

most non-constraints can afford that (as Goldratt noted in *The Goal* [Goldratt and Cox, 1984; 1993])

A *dummy constraint* is a resource constraint that is *inexpensively eliminated*. Pass and Ronen (2003) note two common dummy constraints in marketing and sales: (1) Shortage of inexpensive administrative assistance and (2) lack of laptops and communications equipment such as portable fax machines. They further note three common dummy constraints in R&D: (1) Shortage of low cost components and accessories, (2) shortage of low cost administrative assistance, and (3) lack of computers and IT tools. Breaking these dummy constraints may give a significant elevation to the market constraint.

Smith et al. (1999) also mention DBR as an aid in product development at Allied Signal and Alcoa.

Re-Entrant Flows

Wu and Yeh (2006) describe the use of DBR in a situation in which a part passes through the constraint twice in flowing through the plant, known as “re-entrant flows.” This situation commonly occurs in semiconductor manufacturing. According to Wu and Yeh, the method of scheduling using DBR as described in *The Haystack Syndrome* (Goldratt, 1990) cannot effectively schedule environments with bottleneck re-entrant flows. They cite a number of articles describing the use of DBR in re-entrant flows including Huang et al. (2002), Kayton et al. (1996; 1997), Kim et al. (2003b), Klusewitz and Rerick (1996), Levison (1998), Mosely et al. (1998), Murphy (1994), Murphy and Dederer (1996), Rose et al. (1995a, b), Tyan et al. (2002), and Villforth (1994). Wu and Yeh (2006) then propose a scheduling method for DBR that they feel is appropriate for manufacturing facilities with bottleneck re-entrant flows. Rippenhagen and Krishnaswamy (1998) simulated a wafer

fabrication facility with re-entrant flows using a variety of dispatching rules and Theory of Constraints. Kim et al. (2009) report on a simulation study of a hypothetical wafer facility with re-entrant flows and protective capacity. They are interested in, among other things, the trade-off between protective capacity and protective inventory. The study is based on a six-station line with re-entrant flows and times per part ranging from 8 min to 12 min and protective capacity ranging from 1 min per part to 4 min. They found that simply knowing the percentage of downtime at non-constraints was not sufficient to understand the need for protective capacity and inventory. Specifically, they found that infrequent long outages required more protective capacity/inventory than did frequent short outages even though the proportion of time the station was out was the same. They also found that resource downtime had more impact on the constraint than did processing time variation. They found that allocation of protective capacity throughout the line was more important than protective inventory. WIP inventory involves a tradeoff between Throughput level and cycle time. Beyond some point, adding more inventory does not improve Throughput, so an appropriate level must be chosen.

Recoverable Manufacturing and Remanufacturing

Guide (1997) discusses the successful application of DBR to *recoverable manufacturing*, where *used products are returned from the consumer to the manufacturer, who then remanufactures the product*. Guide uses the term “recoverable product environment” to describe the processes to recover materials via recycling at the end of the product life. Guide (1996) showed that DBR could be a successful production planning and control system for remanufacturing.

Buffer Management Literature

While few of the above simulation or case studies above recognize Buffer Management as a necessary condition for an effective Drum-Buffer-Rope planning and control system, most TOC experts today agree on its vital importance both in expediting orders before they become late and also as the foundation of a process of ongoing improvement. The *TOCICO Dictionary* (Sullivan et al., 2007, 7) defines *buffer management* as “A feedback mechanism used during the execution phase of operations, distribution and project management that provides a means to prioritize work, to know when to expedite, to identify where protective capacity is insufficient, and to resize buffers when needed.” (© TOCICO 2007, used by permission, all rights reserved.)

When an item is released to the floor, it is released into a buffer—constraint, shipping, or assembly buffer depending on the shop’s configuration. Buffers are sized so that each batch or order should arrive at the buffer in time to maintain the buffer approximately half-full. The buffer is actually divided into three regions⁹, each representing one-third of the buffer length. *Region I (Red)* consists of the oldest batches, which should be processed soon; *Region II (Yellow)* represents intermediate batches, about half of which should be in the buffer; and *Region III (Green)* represents the most recently released material, which is generally expected to still be en route to the buffer. If material released into the shop and under the control of the buffer has not yet reached the buffer it is called a *buffer hole*. Simatupang (2000) provides a good description of Buffer Management activities.

There is a person called the buffer manager whose responsibility is to steer material into the buffer on time. Holes in the Green Region of the buffer require no action. If

a hole moves into the Yellow Region, the buffer manager will locate the item and remind the workstation holding the batch that it is soon due in the buffer. If a hole reaches the Red Region, the buffer manager will expedite the batch through the station holding the material and any stations between the batch's location and the physical buffer. The buffer manager will also note the location of the expedited batch and the reason for delay for prioritizing future improvement efforts. Gardiner et al. (1992) state that 90 percent of orders should require no expediting if the buffer is properly sized. The buffer size is dynamic—if too much expediting is occurring, the buffer can be made bigger; if virtually no expediting is occurring, the buffer can be made smaller. Because of Buffer Management and dynamic buffer sizing, the initial size of the buffer is not that critical—if it is initially the wrong size, Buffer Management activities will quickly reveal that fact and the buffer can be resized.

When a job is released to the shop floor, its paperwork should show the due date of the job in the buffer toward which it is moving. The supervisor of each workstation can use this information as an aid in sequencing jobs. The buffer manager of the buffer involved has a sequenced list of jobs due in the buffer, which he can use in determining the location of holes in the buffer and to decide whether to begin investigative action or expediting.

Tseng and Wu (2006) describe Buffer Management in a modified system employing five buffer regions rather than three: early arrival zone, ignored zone, mentioned zone, expediting zone, and delayed zone. The three middle zones correspond to the normal three regions of the buffer, while the first zone represents material released to the shop too early and the fifth zone represents material not processed in time. Simatupang (2000) describes how Buffer Management can be used to direct the application of preventive maintenance activities.

Schragenheim and Dettmer (2001) describe a variation on Buffer Management called the *“red-line control mechanism,”* which *collects data on jobs that are about to be late and assists managers in determining the stability of the shop floor.*

Buffer Sizing

One of the questions that must be addressed in establishing a DBR system is, “What size should the buffers be?” Goldratt has suggested to Jonah courses that an initial buffer size can be developed by taking one-half the current lead time and dividing that time between the constraint time buffer and the shipping buffer. This initial buffer size can be adjusted up or down by whether too few or too many jobs require some expediting via Buffer Management. This suggestion has worked its way into the literature.

Louw and Page (2004) state that the determination of the time buffer lengths is a trial-and-error approach that consists of first determining the initial size of the time buffers through simple empirical rules (Srikanth and Umble, 1997; Tu and Li, 1998). The buffer lengths are then monitored and adjusted through a process known as Buffer Management (Goldratt, 1990; Schragenheim and Ronen, 1991). Goldratt (1990) suggests determining the initial buffer lengths by estimating the current average lead time of the tasks to the specific buffer origin and dividing it by five. Srikanth and Umble (1997) suggest the total time buffer for any product should be approximately one-half the firm’s current manufacturing lead time, whereas Schragenheim and Ronen (1990) suggest a constraint buffer size of three times the minimum cumulative processing time to the constraint.

Louw and Page (2004) use a procedure for estimating the sizes of the time buffers based on a queuing model in a

multi-product open queuing network. Details of this network are beyond the scope of this chapter. Ye and Han (2008) use a mathematical approach to estimate both the time buffer and the assembly buffer sizes.

Weiss (1999) presents a queuing network using separated continuous linear programs, which he says is similar to DBR in that it tends to form buffers at the busiest stations.

Taylor (2002) points out that attempting to remove all system variability is not cost effective. It is better to buffer the constraint and to some extent buffer CCRs in order to protect them from starvation. Taylor simulated MRP, JIT, and DBR systems and compared their influence on a number of operations performance measures.

Some companies have been hesitant to start DBR because they do not know how to set the buffer sizes. Because adjustments to buffer size suggested by Buffer Management will quickly correct any initial buffer size estimate, companies should simply pick a conservative buffer size and get started.

Buffer Sizing and Lead Time

In a serial line with a single resource constraint, there should be two buffers—the time buffer at the constraint and the shipping buffer. The manufacturing lead time through the system should approximate the sum of the two buffer sizes. Even in arrangements that are more complex, this statement would be true unless there is a non-constraint assembly between constraint parts and non-constraint parts and one of the non-constraint parts had a longer lead time to the assembly point than the constraint part. In this case, the lead time should approximate the sum of the assembly buffer and the time allowed to flow from the assembly point to the shipping dock. This relationship, and its importance,

is explained at length in [Chapters 9](#) and [10](#) of this book and by Fry et al. (1991).

TOC and Distribution

Little has been written about the TOC solution to a distribution environment such as a supply chain. However, in the early 1990s, Goldratt utilized a distribution simulator to teach the TOC approach to distribution in various classes. Recently, Schragenheim et al. (2009) published a chapter on the distribution environment with a very thorough treatment. Imagine an environment in which a manufacturer produces a variety of products that are distributed via a set of warehouses to a larger set of retailers. Under traditional management, it is common for the retailer to order an entire season's supply of an item to arrive before the season begins based on a forecast of what sales may be. However, forecasts are always wrong, so the retailer usually runs out of stock before the season ends or has excess stock left at the end of the season that must be sold at greatly reduced prices. In addition, there is the problem of storing that inventory during the season. The TOC solution, as described by Schragenheim et al. (2009), begins with a plan to deliver rather frequently during the season an amount equal to actual sales during the previous delivery period. This requires the retailer to begin the season (and each replenishment cycle) with a stock equal only to the maximum likely sales during the replenishment period. The regional warehouses will hold some stock but most stock will be held in a central warehouse at the manufacturer. This approach takes advantage of the fact that relative variation is much smaller at the manufacturer than it is at the typical retailer. There is less stock in the system, but availability of the item at the retailer is increased because of the frequent deliveries. This is essentially a DBR process applied over the

supply chain. Experience has shown that the increase in Throughput far surpasses any increase in the transportation costs for more frequent deliveries.

Supply Chain Management

Simatupang et al. (2004) discuss the application of TOC to supply chain management. A supply chain consists of different firms that deliver products and services from raw materials to end customers. All the different players such as manufacturers, distributors, and retailers play significant roles in creating value for the ultimate customer. They also note that reliable global performance measures help the chain members to measure progress. They introduce the performance measures *Throughput Dollar Days (TDD), a measure of things done too late and thus endangering Throughput*, and *Inventory Dollar Days (IDD), a measure of things done too early (or that should not have been done) and thus incurring extra inventory carrying costs*.

An important aspect of supply chain management is the decision of whether to out-source particular components—the make-or-buy decision. Of course, component quality is one of the most important aspects of this decision, if not the single most important aspect. Traditionally, cost has been the second most important factor—the cost to make versus the cost to buy. However, in TOC, the decision's impact on Throughput is important and Throughput is impacted in different ways depending on whether making the component requires time at a resource constraint (and perhaps also whether it requires time at a CCR). If making the part requires only non-constraint time and no worker will be laid off because of outsourcing, then traditional cost accounting overestimates the marginal cost of making the part. If the part does require constraint time, then purchasing the part allows additional units of the least

profitable part to be added to the drum schedule, thus increasing Throughput. Traditional cost accounting underestimates the opportunity cost of making the part. Either way, TOC arrives at different numbers for the decision than does traditional accounting. This decision is discussed at length in Gardiner and Blackstone (1991) and is updated by Balakrishnan and Cheng (2005), who point out that if the part is a strategic part, then the cost to buy may not be the most important consideration. The make-or-buy decision is also mentioned in Hilmola (2001).

Walker (2002) provides an excellent discussion of the application of DBR to a supply chain. He discusses how to choose which partner should be the drum, how to tie the rope, total system Inventory measures, and managing as demand goes up and down. Walker (2005) states that the applicability of DBR has been expanded to include the entire supply chain network.

Cox and Walker (2006) have published a board game that uses poker chips in a stochastic supply chain. The players can alter the order policies and batching policies at various points in the supply chain and observe directly the impact on Inventory and service levels.

Service Environment

One of the reasons for keeping buffers as small as practical while not starving the constraint is that if there is too much work in a facility, then workers have a tendency to move back-and-forth between jobs, thus wasting some of their time with extra setups. In [Chapter 21](#) of this handbook, Herman and Goldratt point out that this problem is also true in sales. They include a Current Reality Tree (CRT) describing the problem. Umble and Umble (2006) describe how Buffer Management was used in two accident and emergency facilities in Oxfordshire, UK to track patient care. Motwani,

Klein, and Harowitz (1996a; 1996b) have a two-part article describing the use of TOC and DBR in services, in general, with a specific example from health care.

TOC and Other Modern Philosophies[10](#)

TOC and Lean

Dettmer (undated) states that the Toyota Production System is better known than TOC primarily because it is a much older system (development started in the 1950s) versus TOC in the 1980s). He continues by saying that both systems use continuous improvement and have the goal of obtaining higher profits. Both methods recognize that the customer is the final arbiter of what *value* is.

Berry and Smith (2005) provide a comparison of TOC with Lean and with several other approaches—MRP, MRP II, ERP, and Supply Chain Management.

Sale and Inman (2003) surveyed over 900 firms and received 93 responses. They found that firms using TOC had significantly higher performance improvement than firms using JIT and traditional manufacturing.

Moore and Scheinkopf (1998) compare TOC and Lean. Both TOC and Lean concentrate on continuous improvement and control the flow of material on the shop floor. Both have had dramatic improvements of profitability and lead times and have resulted in operations being drastically simplified.

TOC and TQM

Lepore and Cohen (1999) suggest there are many synergies between TOC and Total Quality Management (TQM). Cohen is one of Goldratt's early partners. Lepore is an academic specializing in Total Quality Management. They suggest a

10-step strategy for implementing the two philosophies together. Step 4 is to implement the 5FS. Step 5 is to implement Buffer Management. However, the book contains little specifics on DBR, per se.

TOC, Lean, and Six Sigma

Pirasteh and Farah (2006, 32-33) state that the top elements of TOC, Lean and Six Sigma work well together. They report on a company that combined the “best components” of these three approaches into what they called TLS. They applied Six Sigma alone to 11 plants, Lean alone to 4 plants, and TLS to 6 plants. They measured plant performance regarding “on-time delivery, warranty costs, customer returns, Inventory reduction, cycle time reduction, and scrap expense.” The company concluded that “the TLS process improvement methodology delivered considerably higher cost savings to the company.”

Problems with DBR

One of the most frequently mentioned conceptual problems with DBR is the issue of wandering bottlenecks, that is, frequent changes in the resource constraint station. The usual cause for this is given that in the case of job shops, shifting of the order mix causes such shifts in the most overloaded resource. Goldratt disputes this issue as a problem out of both his experience and using the logic that even in a job shop there is usually one primary machine or skill that is driving the bulk of orders. An occasional shift in bottlenecks is not a problem as the shop can change its focus occasionally. Hurley and Kadipasaoglu (1998) speculate on the causes of wandering bottlenecks. They demonstrate that changing product mix is a minor contributor to this problem and that the primary cause is

management actions in response to inappropriate performance measures. One such policy is the continuing use of non-bottleneck utilization as a performance measure—seeing unused protective capacity as a waste. Releasing material faster than the rope requires builds up inventory at many stations and disrupts the drum schedule with unneeded work delaying needed work. Increasing batch sizes to minimize setups can lead to large jobs at non-bottlenecks clogging the shop and creating an unnecessary shift of the bottleneck. They conclude that only in a small number of cases is a product-mix-driven wandering bottleneck truly an issue.

Riezebos, Korte, and Land (2003) report on a problem with maintaining lead times in a DBR implementation, which they corrected using Workload Control to better manage the release of material into the shop, thereby maintaining an appropriate buffer size.

Simons et al. (1996) discuss the difficulty of scheduling a DBR managed system with multiple CCRs where they correctly stipulate that a CCR need not be a bottleneck. They follow the process for applying DBR as outlined by Goldratt (1990, 241-3) in *The Haystack Syndrome*. They created a “diverse set of benchmark problems” on which to test the efficacy of the general DBR algorithm. They used a branch-and-bound approach to obtain optimal schedules and found that in the presence of multiple CCRs, the DBR solution averaged within 3 percent of optimal.

Floating or Multiple Bottlenecks

A situation that is frequently proposed to give DBR problems is the existence of multiple or floating bottlenecks; that is, bottlenecks that change over time because of seasonal or long-run changes in the product mix. Lawrence and Buss (1994) state that balanced utilization rates increase the

shifting bottleneck problem. They further state that increasing capacity at non-bottleneck work centers is the “best hope” for improving shop performance.

Simons and Simpson (1997) defend DBR’s ability to contend with multiple constraints. The Goal System utilizes an iterative procedure to schedule multiple constraints to “accommodate interaction.”

Guan et al. (2007) report simulating an electronics manufacturing system with multiple bottlenecks. Lenort and Samolejova (2007) report on identifying floating bottlenecks in metallurgical production and using such identification to maximize output.

Summary and Conclusions

This literature amply demonstrates that DBR is an effective and efficient system for planning and control of both manufacturing and service organizations. It has been applied successfully in a wide range of organizations. Reported problems are few and seem to occur primarily where implementers develop inadequate protective capacity. Two remaining issues for research are the ideal levels of protective capacity and the correct initial buffer sizes. Buffer sizing is a short-term problem because buffer sizes can be adjusted quickly using information developed through Buffer Management. Protective capacity usually cannot be established precisely as a manager might hope, as equipment is available only in certain sizes so a piece of equipment yielding the desired amount of protective capacity may not exist. The most recent application of DBR is to supply chains. Early papers on this issue have argued that application of the 5FS and DBR within supply chains is both possible and beneficial.

For future research on cases, it would be helpful if the investigators would be specific about whether the plant was

a V, A, or T configuration; whether the market or internal resource was the constraint; how buffers were sized initially; and how Buffer Management was achieved. This information would enable the reader to understand the implementation more thoroughly.

Future simulation research on more complex plants regarding protective capacity and protective inventory is needed. Most of such simulation research to date has involved only a few stations in an I formation.

References

- Amen, M. 2000. "Heuristic methods for cost-oriented assembly line balancing: A survey," *International Journal of Production Economics* 68:1-14.
- Amen, M. 2001. "Heuristic methods for cost-oriented assembly line balancing: A comparison on solution quality and computing time," *International Journal of Production Economics* 69:255-264.
- Andrews, C. and Becker, S. W. 1992. "Alkco Lighting and its journey to Goldratt's goal," *Total Quality Management* 3:71-95.
- Atwater, J. B. 1991. The impact of protective capacity on the output of a typical unblocked flow shop. Doctoral diss., University of Georgia.
- Atwater, J. B. and Chakravorty, S. 1994. "Does protective capacity assist managers in competing along time-based dimensions?" *Production and Inventory Management Journal* 35:53-59.

- Atwater, J. B. and Chakravorty, S. S. 1996. "The impact of restricting the flow of inventory in serial production systems." *International Journal of Production Research* 34:2657-2669.
- Atwater, J. B. and Chakravorty, S. 2002. "A study of the utilization of capacity constrained resources in drum-buffer-rope systems," *Production and Operations Management* 12:259-273.
- Atwater, J. B., Stephens, A. A., and Chakravorty, S. S. 2004. "Impact of scheduling free goods on the throughput performance of a manufacturing operation," *International Journal of Production Research* 42:4849-4869.
- Balakrishnan, J. and Cheng, C. H. 2005. "The Theory of Constraints and the make-or-buy decision—An update and review," *Journal of Supply Chain Management* 41:40-47.
- Becker, C. and Scholl, A. 2006. "A survey on problems and methods in generalized assembly line balancing," *European Journal of Operational Research* 168:694-715.
- Belvedere, V. and Grando, A. 2005. "Implementing a pull system in batch/mix process industry through Theory of Constraints: A case study," *Human Systems Management* 24:3-12.
- Berry, R. and Smith, L. B. 2005. "Conceptual foundations for the Theory of Constraints," *Human Systems Management* 24:83-94.
- Betterton, C. E. and Cox III, J. F. 2009. "Espoused drum-buffer-rope flow control in serial lines: A comparative

study of simulation models," *International Journal of Production Economics* 117(1):66-79.

Betz, H. J. 1996. "Common sense manufacturing: A method of production control," *Production and Inventory Management Journal* 37:77-81.

Blackstone, J. H., Jr. and Cox III, J. F. 2002. "Designing unbalanced lines—Understanding protective capacity and protective inventory," *Production Planning & Control* 13:416-423.

Boorstin, D. 1965. *The Americans: The National Experience*. New York: Random House.

Boyd, L. and Gupta, M. 2004. "Constraints management: What is the theory?" *International Journal of Operations and Production Management* 24:350-371.

Chakravorty, S. S. 1996. "Robert Bowden, Inc.: A case study of cellular manufacturing and drum-buffer-rope implementation," *Production and Inventory Management Journal* 37(3): 15-19.

Chakravorty, S. S. 2000. "Improving a V-plant operation: A window manufacturing case study," *Production and Inventory Management Journal* 41(3):37-42.

Chakravorty, S. S. and Atwater, J. B. 1996. "A comparative study of line design approaches for serial production systems," *International Journal of Operations and Production Management* 16(6):91-108.

Chakravorty, S. S. and Atwater, J. B. 2005. "The impact of free goods on the performance of drum-buffer-rope

scheduling systems," *International Journal of Production Economics* 95:347-357.

Coman, A., Koller, G., and Ronen, B. 1996. "The application of focused management in the electronics industry," *Production and Inventory Management Journal* 37(2):63-70.

Conti, R. F. and Warner, M. 1997. "Technology, culture and craft: Job tasks and quality realities," *New Technology Work and Employment* 12:123-135.

Conway, R. 1997. "Comments on an exposition of multiple constraint scheduling," *Production and Operations Management* 6:23-24.

Corbett, T. and Csillag, J. 2001. "Analysis of the effects of seven drum-buffer-rope implementations," *Production and Inventory Management Journal* 42(3):17-23.

Cox III, J. F., III and Spencer, M. 1998. *The Constraints Management Handbook*. Boca Raton, FL: St. Lucie Press.

Cox III, J. F., III and Walker, E. D., II. 2006. "The poker chip game: A multiproduct, multi-echelon, stochastic supply chain network useful for teaching the impacts of pull versus push inventory policies on link and chain performance," *Inform Transactions on Education* 6(3):3-19.

Danos, G. 1996. "Dixie reengineers scheduling—And increases profit 300 percent," *APICS The Performance Advantage* 3:28-31.

Demmy, W. S. and Demmy, B. S. 1994. "Drum-buffer-rope scheduling and pictures for the yearbook,"

Production and Inventory Management Journal
35(3):45-47.

Demmy, W. S. and Petrini, A. B. 1992. "The Theory of Constraints: A new weapon for depot maintenance," *Air Force Journal of Logistics*, 16(3):6-10.

Dettmer, H. W. Undated. *Beyond Lean Manufacturing: Combining Lean and the Theory of Constraints for Higher Performance*. Port Angeles, WA: Goal Systems International.

Duclos, L. K. and Spencer, M. S. 1995. "The impact of a constraint buffer in a flow shop," *International Journal of Production Economics* 42:175-185.

Fawcett, S. and Peterson, J. 1991. "Understanding and applying constraint management in today's manufacturing environments," *Production and Inventory Management Journal* 32(3):46-55.

Finch, B. J. and Luebbe, R. 1995. "The impact of learning rate and constraints on production line performance," *International Journal of Production Research* 33:631-642.

Frazier, G. and Reyes, P. 2000. "Applying synchronous manufacturing concepts to improve production performance in high-tech manufacturing," *Production and Inventory Management Journal* 41(3):60-65.

Fry, T. D. 1990. "Controlling input: The real key to shorter lead times," *The International Journal of Logistics Management* 1:1-12.

Fry, T. D., Karan, K. R., and Steele, D. C. 1991. "Implementing drum-buffer-rope to control

manufacturing lead time," *The International Journal of Logistics Management* 2:12-18.

Gardiner, S. C. and Blackstone Jr., J. H. 1991. "The Theory of Constraints and the make-or-buy decision," *International Journal of Purchasing and Materials Management* 27(3):38-43.

Gardiner, S. C., Blackstone, J., and Gardiner, L. 1992. "Drum-buffer-rope and buffer management: Impact on production management study and practices," *International Journal of Operations and Production Management* 13(6):68-78.

Gardiner, S. C., Blackstone, J., and Gardiner, L. 1994. "The evolution of the Theory of Constraints," *Industrial Management* May/June:13-16.

Goldratt, E. M. 1988. "Computerized shop floor scheduling," *International Journal of Production Research* 26:443-455.

Goldratt, E. M. 1990. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Crotonon-Hudson, NY: North River Press.

Goldratt, E. M. 2003a. *My Saga to Improve Production in Production The TOC Way*. rev. ed. Great Barrington, MA: North River Press.

Goldratt, E. M. 2003b. *Production: The TOC Way*. rev. ed. Great Barrington, MA: North River Press.

Goldratt, E. M. 2009. "Standing on the shoulders of giants," *The Manufacturer* June. <http://www.themanufacturer.com/uk/content/9280/Sta>

nding_on_the_shoulders_of_giants. (accessed February 4, 2010).

Goldratt, E. M. and Cox, J. 1984, 1993. *The Goal*. Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. and Fox, R. E. 1986. *The Race*. Croton-on-Hudson, NY: North River Press.

Grosfeld-Nir, A. and Ronen, B. 1992. "A single-bottleneck system with binomial yields and rigid demand," *Management Science* 39:650-654.

Guan, Z. L., Peng, Y. F., Zeng, X. L., and Shao, X. Y. 2007. "TOC/DBR based production planning and control in a manufacturing system with multiple constraints," *Industrial Engineering and Engineering Management* 1078-1082.

Guide, V. D. 1995. "A simulation model of drum-buffer-rope for production planning and control at a naval aviation depot," *Simulation* 65:157-168.

Guide, V. D. 1996. "Scheduling using drum-buffer-rope in a remanufacturing environment," *International Journal of Production Research* 34:1081-1091.

Guide, V. D. 1997. "Scheduling with priority dispatching rules and drum-buffer-rope in a recoverable manufacturing system," *International Journal of Production Economics* 53:101-116.

Guide, V. D. and Ghiselli, G. 1995. "Implementation of drum-buffer-rope at a military rework depot engine works," *Production and Inventory Management Journal* 36(3):79-83.

- Gupta, M. 2003. "Constraints management—Recent advances and practices," *International Journal of Production Research* 41:647-659.
- Gupta, M., Ko, H. J., and Min, H. 2002. "TOC-based performance measures and five focusing steps in a job-shop manufacturing environment," *International Journal of Production Research* 40:907-930.
- Gupta, S. 1997." Supply chain management in complex manufacturing", *IIE Solutions* March:18-23.
- Hall, R. W. 1997. "Just-in-time concepts: Scope and applications," in Greene, J. Editor. *Production & Inventory Control Handbook*. New York: McGraw Hill.
- Hasgul, S. and Kartal, K. 2007. Analyzing a drum-buffer-roppe scheduling system executability through simulation. Working paper. Eskisehir Osmangazi University, SCSC, 1243-1249.
- Heizer, J. H. 1998. "Determining responsibility for development of the moving assembly line." *Journal of Management History* 4:94-103.
- Hilmola, O. P. 2001. "Theory of Constraints and outsourcing decisions," *International Journal of Manufacturing Technology and Management* 3:517-527.
- Huang, J. Y. and Sha, D. Y. 1998. "Constructing procedures of an effective production activity control technique for a wafer fabrication environment," *International Journal of Industrial Engineers* 5:235-243.

- Huang, S. H., Dismukes, J. P., Shi, J., Wang, Q. S. G., Razzak, M. A., and Robinson, D. E. 2002. "Manufacturing system modeling for productivity improvement," *Journal of Manufacturing Systems* 21:249-259.
- Huff, P. 2001. "Using drum-buffer-rope scheduling rather than just-in-time production," *Management Accounting Quarterly* Winter:36-40.
- Hurley, S. F. and Kadipasaoglu, S. 1998. "Wandering bottlenecks: Speculating on true causes," *Production and Inventory Management Journal* 39(4):1-4.
- Hurley, S. F. and Whybark, D. C. 1999. "Inventory and capacity trade-offs in a manufacturing cell," *International Journal of Production Economics* 59:203-212.
- Jackson, G. C. and Low, J. T. 1993. "Constraint management: A description and assessment." *The International Journal of Logistics Management* 4(2):41-48.
- Kadipasaoglu, S. N., Xiang, W., Hurley, S. F., and Khumwala, B. M. 2000. "A study on the extent and location of protective capacity in flow systems," *International Journal of Production Economics* 63:217-228.
- Kayton, D., Tayner, T., Schwartz, C., and Uzsoy, R. 1996. "Effects of dispatching and down time on the performance of wafer fabs operating under theory of constraints," *1996 International Electronics Manufacturing Technology Symposium*. Austin, TX, 49-56.

- Kayton, D., Tayner, T., Schwartz, C., and Uzsoy, R. 1997. "Focusing maintenance improvement efforts in a wafer fabrication facility operating under the theory of constraints," *Production and Inventory Management Journal* 38(4):51-57.
- Kim, S., Cox III, J. F., and Mabin, V. J. 2009. "An exploratory study of protective inventory in a re-entrant line with protective capacity," *International Journal of Production Research*. Currently available online at <http://dx.doi.org/10.1080/00207540902991666>.
- Kim, S., Davis, K. R., and Cox III, J. F. 2003a. "An investigation of output flow control, bottleneck flow control and dynamic flow control mechanisms in various simple lines scenarios," *Production Planning and Control* 14:15-32.
- Kim, S. S., Davis, K. R., and Cox III, J. F. 2003b. "Investigation of flow mechanisms in semiconductor wafer fabrication," *International Journal of Production Research* 41:681-698.
- Klusewitz, G. and Rerick, R. 1996. "Constraint management through the drum-buffer-rope system," Cambridge, MA. 1996 *IEEE/SEMI Advanced Semiconductor Manufacturing Conference*. Cambridge, MA, 7-12.
- Kosturiak, J. and Gregor, M. 1998. "FMS simulation: Some experience and recommendations," *Simulation Practice and Theory* 6:423-442.
- Koziol, D. 1988. "How the constraint theory improved a job-shop operation," *Management Accounting* 69(11):44-49.

- Lambrecht, M. and Alain, S. 1990. "Buffer stock allocation in serial and assembly type of production lines," *International Journal of Operations and Production Management* 10(2):47-61.
- Lambrecht, M. and Decaluwe, L. 1988. "JIT and constraint theory: The issue of bottleneck management," *Production and Inventory Management Journal* 29(3):61-65.
- Lawrence, S. R. and Buss, A. H. 1994. "Shifting production bottlenecks: Causes, cures, and conundrums," *Production and Operations Management* 3:21-37.
- Lea, B. R. and Min, H. 2003. "Selection of management accounting systems in just-in-time and theory of constraints-based manufacturing," *International Journal of Production Research* 41:2879-2910.
- Lenort, R. and Samolejova, A. 2007. "Analysis and identification of floating capacity bottlenecks in metallurgical production," *Metalurgica* 46:61-66.
- Lepore, D. and Cohen, O. 1999. *Deming and Goldratt*. Great Barrington, MA: North River Press.
- Levison, W. A. 1998. *Leading the Way to Competitive Excellence: The Harris Mountaintop Case Study*. Milwaukee, WI: American Society for Quality.
- Lindsay, C. G. 2005. "TOC in the DC," *Industrial Engineer* 37(6):29-33.
- Louw, L. and Page, D. C. 2004. "Queuing network analysis approach for estimating the sizes of the time buffers in theory of constraints-controlled production

systems," *International Journal of Production Research* 42:1207-1226.

Luck, G. 2004. "New market innovation through supply chain management," *Critical Eye Magazine* March-May. 42-45. Available at [http://ashridgemanagementcollege.net/Website/IC.nsf/wFARATT/New Market Innovation Through Supply Chain](http://ashridgemanagementcollege.net/Website/IC.nsf/wFARATT/New%20Market%20Innovation%20Through%20Supply%20Chain).

Mabin, Vee. J. and Balderstone, S. J. 2000. *The World of the Theory of Constraints*. Boca Raton, FL: St. Lucie Press.

Mabin, V. J. and Balderstone, S. J. 2003. "The performance of the theory of constraints methodology: Analysis and discussion of successful TOC applications," *International Journal of Operations & Production Management* 23:568-595.

Mabin, V. and Davies J. 1999. "Reframing the Product Mix Problem Using the Theory of Constraints, in Oregon in the New Millennium," Proceedings of the 34th Annual Conference of ORSNZ, Hamilton, OR, 227-236.

Mabin, V. I. and Davies, J. 2003. "Framework for understanding the complementary nature of TOC frames: Insights from the product mix dilemma," *International Journal of Production Research* 41:661-680.

Moore, R. and Scheinkopf, L. 1998. Theory of constraints and lean manufacturing: Friends or foes. Working paper. Chesapeake Consulting, Inc.

- Mosely, S. A., Teyner, T., and Uzsoy, R. 1998. "Maintenance scheduling and staffing policies in a wafer fabrication facility," *IEEE Transactions in Semiconductor Manufacturing* 11:316-323.
- Motwani, J., Klein, D., and Harowitz, R. 1996a. "The theory of constraints in services: Part 1—The basics," *Managing Service Quality* 6:53-56.
- Motwani, J., Klein, D., and Harowitz, R. 1996b. "The theory of constraints in services: Part 2—Examples from health care," *Managing Service Quality* 6(2):30-34.
- Murphy, R. E. 1994. "Synchronous flow management (SFM) principles in a wafer fabrication facility," *1994 IEEE/SEMI Advanced Semiconductor Manufacturing Conference*. Cambridge, MA, 179-184.
- Murphy Jr., R. E. and Dederer, C. R. 1996. "Holistic TOC for maximum profitability," *Proceedings of the Advanced Semiconductor Manufacturing Conference and Workshop*. IEEE, 242-249.
- Pass, S. and Ronen, B. 2003. "Management by the market constraint in the hi-tech industry," *International Journal of Production Research* 41:713-724.
- Pinedo, M. 1997. "Commentary on 'An exposition of multiple constraint scheduling as implemented in the Goal System,'" *Production and Operations Management Society* 6:25-27.
- Pirasteh, R. M. and Farah, K. S. 2006. "Continuous improvement trio," *APICS Magazine* May:31-33.

- Politou, A. and Georgiadis, P. Undated. *Production Planning and Control in Flow Shop Operations Using Drum-Buffer-Rope Methodology: A System Dynamics Approach*. Thessaloniki, Greece: Aristotle University of Thessaloniki.
- Radovilsky, Z. 1994. "Estimating the size of the time buffer in the Theory of Constraints: Implications for management," *International Journal of Management* 11:839-847.
- Radovilsky, Z. D. 1998. "A quantitative approach to estimate the size of the time buffer in the theory of constraints," *International Journal of Production Economics* 55:113-119.
- Rahman, S.-ur. 1998. "Theory of Constraints: A review of the philosophy and its applications," *International Journal of Operations and Production Management* 18:336-355.
- Reimer, G. 1991. "Material requirements planning and Theory of Constraints: Can they coexist? A case study," *Production and Inventory Management Journal* Fourth Quarter:48-52.
- Rerick, R. A. 1997. "Fab 6 pipeline constraint management implementation at Harris Semiconductor Corp." *Microelectronics Journal* 28(2): viii-ix.
- Riezebos, J., Korte, G. J., and Land, M. J. 2003. "Improving a practical DBR buffering approach using Workload Control," *International Journal of Production Research* 41:699-712.
- Rippenhagen, C. and Krishnaswamy, S. 1998. "Implementing the theory of constraints philosophy in

highly reentrant systems," *Proceedings of the 1998 Winter Simulation Conference*, 993-996.

Ronen, B. and Spector, Y. 1992. "Managing system constraints: A cost/utilization approach." *International Journal of Production Research* 30:2045-2061.

Rose, E., Odom, R., Dunbar, R., and Hinchman, J. 1995a. "How TOC and TPM together to build the quality toolbox of SDWTs," *1995 IEEE CMPT International Manufacturing Technology Symposium*, 56-59.

Rose, E., Odom, R., Murphy, R., and Behnke, L. 1995b. "SDWT requires tools to be successful," *1995 IEEE/SEMI Advanced Manufacturing Conference*, 327-332.

Russell, G. R. and Fry, T. D. 1997. "Order review/release and lot splitting in drum-buffer-rope," *International Journal of Production Research* 35:827-845.

Sale, M. L. and Inman, R. A. 2003. "Survey-based comparison of performance and change in performance of firms using traditional manufacturing, JIT and TOC," *International Journal of Production Research* 41:829-844.

Schaeffers, J., Aggounne, R., Becker, F., and Fabri, R. 2004. "TOC-based planning and scheduling model," *International Journal of Production Research* 42:2639-2649.

Schragenheim, E., Cox, J. F., and Ronen, B. 1994. "Process flow industry—Scheduling control using the Theory of Constraints," *International Journal of Production Research* 32:1867-1877.

- Schrageheim, E. and Dettmer, H. W. 2001. *Manufacturing at Warp Speed*. Boca Raton, FL: St. Lucie Press.
- Schrageheim, E., Dettmer, H. W., and Patterson, J. W. 2009. *Supply Chain Management at Warp Speed*. Boca Raton, FL: CRC Press.
- Schrageheim, E. and Ronen, B. 1990. "Drum-buffer-ropes shop floor control," *Production and Inventory Management Journal* 31(3):18-23.
- Schrageheim, E. and Ronen, B. 1991. "Buffer management: A diagnostic tool for production control," *Production and Inventory Management Journal* 32(2):74-79.
- Simatupang, T. M. 2000. Utilization of buffer management to build focused productive maintenance. Working paper. New Zealand: Massey University.
- Simatupang, T. M., Wright, A. C., and Sridharan, R. 2004. "Applying the theory of constraints to supply chain collaboration," *Supply Chain Management: An International Journal* 9:57-70.
- Simons, Jr., J. V. and Simpson III, W. P. 1997. "An exposition of multiple constraint scheduling as implemented in the Goal System (formerly Disaster™)," *Production and Operations Management* 6:3-22.
- Simons, Jr., J. V., Simpson, W. P., Carlson, B. J., James, S. W., Lettiere, C. A., and Mediate, Jr., B. A. 1996. "Formulation and solution of the drum-buffer-rope constraint scheduling problem (DBRCSP),"

International Journal of Production Research 34:2405-2420.

Smith, G. R., Herbein, W. C., and Morris, R. C. 1999. "Front-end innovation at AlliedSignal and Alcoa," *Research-Technology Management* 42(6):15-24.

Sorensen, C. 1956. *My Forty Years with Ford*. New York: W. W. Norton & Company.

Spearman, M. L. 1997. "On the Theory of Constraints and the Goal System," *Production and Operations Management* 6:28-33.

Spencer, M. S. 1991. "Using *The Goal* in an MRP system," *Production and Inventory Management Journal* Fourth Quarter:22-27.

Spencer, M. S. 1994. "Economic theory, cost accounting and theory of constraints: An examination of relationships and problems," *International Journal of Production Research* 32:299-308.

Spencer, M. S. and Cox III, J. F. 1994. "Optimum Production Technology (OPT) and the Theory of Constraints (TOC): Analysis and genealogy," *International Journal of Production Research* 33:1495-1504.

Spencer, M. S. and Cox III, J. F. 1995. "Master production scheduling in a Theory of Constraints environment," *Production and Inventory Management Journal* First Quarter:8-14.

Spencer, M. S. and Wathen, S. 1994. "Applying the Theory of Constraints' process management

technique to an administrative function at Stanley Furniture," *National Productivity Review* July:379-385.

Srikanth, M. L. and Umble, M. M. 1997. *Synchronous Management: Profit Based Manufacturing for the 21st Century*, volume 1. Guilford: Spectrum, 235-298.

Steele, D. C., Philipoom, P. R., Malhotra, M. K., and Fry, T. D. 2005. "Comparisons between drum-buffer-rope and material requirements planning: A case study," *International Journal of Production Research* 48:3181-3208.

Stein, R. E. 1996. *Reengineering the Manufacturing System: Applying the Theory of Constraints (TOC)*. New York: Marcel Dekker.

Sugimori, Y., Kusunoki, K., Cho, F. and Uchikawa, S. 1977. "Toyota Production System and Kanban system: Materialization of just-in-time and respect-for-human system," *International Journal of Production Research* 15:(6):553-564.

Sullivan, T. T., Reid, R. A., and Cartier, B. Editors. 2007. *The TOCICO Dictionary*. Theory of Constraints International Certification Organization. Available online at <http://www.tocico.org/?page=dictionary>.

Taylor III, L. J. 1999. "A simulation study of work-in-process inventory drive systems and their effect on financial measures," *Integrated Manufacturing Systems* 10:306-315.

Taylor III, L. J. 2000. "A simulation study of work-in-process inventory drive systems and their effect on operational measures," *British Journal of Management* 11:47-59.

- Trietsch, D. 2005. "From management by constraints (MBC) to management by criticalities (MBC II)," *Human Systems Management* 24:105-115.
- Tseng, M. E. and Wu, H. H. 2006. "The study of an easy-to-use DBR and BM system," *International Journal of Production Research* 44:1449-1478.
- Tu, Y. M. and Li, R. K. 1998. "Constraint time buffer determination model," *International Journal of Production Research* 36:1091-1103.
- Tyan, J. C., Chen, J. C., and Wang, F. K. 2002. "Development of a state-dependent dispatch rule using theory of constraints in near-real-world wafer fabrication," *Production Planning & Control* 13:253-261.
- Umble, M. and Srikanth, M. L. 1995. *Synchronous Manufacturing: Principles for World Class Excellence*. Wallingford, CT: Spectrum.
- Umble, M. M. and Umble, E. J. 1999. "Drum-buffer-rope for lower inventory," *Industrial Management* September-October:24-33.
- Umble, M., Umble, E., and Von Deylan, L. 2001. "Integrating Enterprise Resources Planning and Theory of Constraints: A case study," *Production and Inventory Management Journal* 42(2):43-48.
- Umble, M. M. and Umble, E. J. 2006. "Utilizing buffer management to improve performance in a healthcare environment," *European Journal of Operational Research* 174:1060-1075.

- Umble, M., Umble, E., and Murakami, S. 2006. "Implementing theory of constraints in a traditional Japanese manufacturing environment: The case of Hitachi Tool Engineering," *International Journal of Production Research* 44:1863-1880.
- Vaidyanathan, B. S., Miller, D. M., and Park, Y. H. 1998. "Application of discrete event simulation in production scheduling," *Proceedings of the 1998 Winter Simulation Conference* 2:965-971.
- Vermaak, W. and Ventner, D. Undated. "Using simulation and the Theory of Constraints to optimize materials handling systems." [http://login.totalweblite.com/Clients/doublearrow/belton2001/5.using simulation and the theory of constraints to optimize materials handling systems.pdf](http://login.totalweblite.com/Clients/doublearrow/belton2001/5.using%20simulation%20and%20the%20theory%20of%20constraints%20to%20optimize%20materials%20handling%20systems.pdf).
- Villforth, R. 1994. "Applying constraint management theory in a wafer fab," *1994 IEEE/SEMI Advanced Semiconductor Manufacturing Conference* 175-178.
- Walker, W. T. 2002. "Practical application of drum-buffer-ropes to synchronize a two-stage supply chain," *Production and Inventory Management Journal* 43(3):13-23.
- Walker, W. T. 2005. "Emerging trends in supply chain architecture," *International Journal of Production Research* 43:3517-3528.
- Watson, K. J., Blackstone Jr., J. H., and Gardiner, S. C. 2007. "The evolution of a management philosophy: The Theory of Constraints," *Journal of Operations Management* 25:387-402.

- Weiss, G. 1999. "Scheduling and control of manufacturing systems—A fluid approach," *Proceedings of the 37 Allerton Conference* 577-586.
- Wolffarth, G. 1998. "Organizational issues and strategies to consider when implementing computerized DBR," April 16-17. *APICS-Constraint Management Symposium Proceedings: Making Common Sense a Common Practice, Seattle, WA*, 25-28.
- Wu, H. H. and Yeh, M. L. 2006. "A DBR scheduling method for manufacturing environments with bottleneck re-entrant flows," *International Journal of Production Research* 44:883-902.
- Wu, S. -Y., Morris, J. S., and Gordon, T. M. 1994. "A simulation analysis of the effectiveness of drum-buffer-rope scheduling in furniture manufacturing," *Computers & Industrial Engineering* 26:757-765.
- Ye, T. and Han, W. 2008. "Determination of buffer sizes for drum-buffer-rope (DBR)-controlled production systems," *International Journal of Production Research* 46:2827-2844.
- Yenradee, P. 1994. "Application of Optimised Production Technologies in a capacity constrained flow shop: A case study in a battery factory," *Computers and Industrial Engineering* 27:217.

About the Author

John H. Blackstone, Jr., is a Professor in the Department of Management at the University of Georgia. He has taught courses in Operations Management, Productivity Management, and Quality Management and Manufacturing

Simulation to graduate and undergraduate students. John was raised in Auburn, Alabama, where he attended primary and secondary schools. After a stint as an accounting clerk in the U.S. Air Force, John attended Auburn University where he received a Bachelor of Science and Master of Science in Economics. John then attended Texas A&M University where he received a Doctor of Philosophy in Industrial Engineering. John began his teaching career in 1979 at Auburn University and transferred to the University of Georgia in 1983. John was introduced to the concepts of Eli Goldratt when he read an article in *Fortune* in the fall of 1983 and attended a lecture by Bob Fox at the APICS Zero Inventory Crusade that same year. When *The Goal* was published in 1984, John began using the book as part of his Operations Management course and continues to use it. John attended a Jonah Course in January 1989, and for about three years helped to teach the course to both practitioners and academics. John has authored or coauthored 40 academic articles, several of which have TOC as a topic. He is especially interested in studying the ideal shape and quantity of protective capacity in various situations. He has also authored or coauthored four academic textbooks. John is married to the former Melissa Swift and has four children and four grandchildren.

CHAPTER 8

DBR, Buffer Management, and VATI Flow Classification

Mokshagundam (Shri) Srikanth

Introduction

The *Theory of Constraints* (TOC) provides a simple and practical approach to the problem of managing complex systems. In this chapter, we discuss the application of TOC to production or manufacturing environments. Production/manufacturing environments are among the most complex of systems, characterized by high levels of dependency and variability. Planning the work of the many resources (often 100 or more), procuring the supply of materials from vendors, and coordinating all of these tasks in such a way as to meet committed delivery dates are truly challenging tasks. The development of computers and computer-based planning systems has been a major facilitator for these challenging tasks. Unfortunately, computers have not been a panacea and, in many ways, the use of computers has aggravated the problem—for example, it has been the author's experience that the nervousness¹ in manufacturing supply chains is higher when the supply chain is managed by a sophisticated Enterprise Resources Planning (ERP) (or material requirements planning [MRP]) system.

In this chapter, we first show the application of the TOC approach to managing production environments—known as

*Drum-Buffer-Rope (DBR) and Buffer Management (BM). DBR and BM are the systems that emerge from the application of the Five Focusing Steps (5FS). DBR is the TOC methodology for planning and BM is the TOC methodology for execution control. The term *planning* is used for those activities that start with known market demand and generate the plans for managing the flow of material through the factory including identifying what purchased materials will be needed and when. Execution control refers to the actions that are taken during the execution phase of the plan developed previously. These actions are necessary to ensure that the plans are followed and include the corrective actions that must be taken when deviations from the plan threaten to compromise delivery dates and Throughput of the system.*

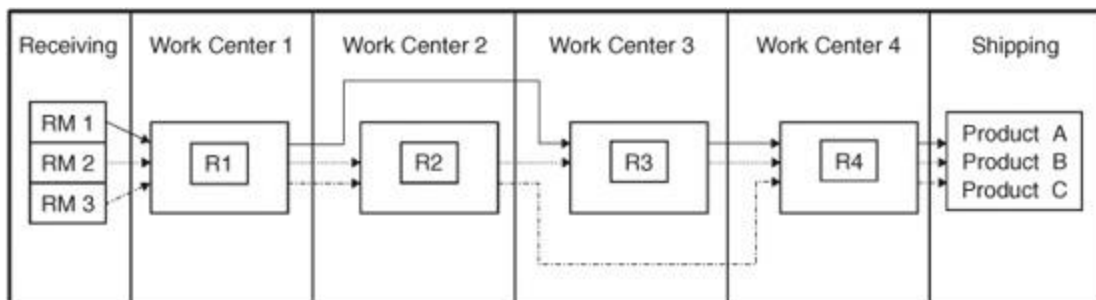


FIGURE 8-1 Resource centric representation of a plant producing three products (A, B, C) with four resources (R1, R2, R3, R4).

After explaining these systems and their logic with simple examples, we next move to a discussion of complex, real-life flows. Real-life production environments are characterized by high levels of detail complexity and high levels of dynamic complexity.² Many of these elements, especially the ones in detail complexity, are specific to the individual environment and make each one appear to be different and unique. However, the behavior of these systems as a whole are characterized more by the way their

dynamic complexity relates one to the other. These relationships and their many apparently different operations exhibit similar behaviors with respect to operational performance as measured by on-time deliveries, system inventories, production lead times, and so on. Third, we present a classification of production operations based on the structure of the product as contained in the bill-of-material and routing or process information. We classify the product flows into four major types—V, A, T, and I—or a combination of these four types. The real power of this classification is that operations that belong to a particular V, A, T, or I type will share similar performance characteristics and business problems to others in the same group. The application of DBR in each type of plant is also discussed.

Managing Flow—Planning and DBR

The Need for a Focus on Flow

A production operation is characterized by a number of resources that typically occupy fixed spots on the factory floor. Materials move from one resource to another in accordance with the rules specified in the routing sheet for the specific material/product. Typically, we think of the factory or production operation from this spatial or static perspective. We will call this a resource centric view of the operation. For the simple case of a factory that has four resources R1, R2, R3, and R4 and makes three products—identified as Product A, Product B, and Product C—the resource centric view of the operation is depicted in [Fig. 8-1](#). The solid black line (—) represents the path in which Product A moves from RM 1 (raw material 1) through the various resources as it is converted from raw material to a finished

product. The dotted line (. . .) represents the path in which Product B moves from RM 2 through the various resources, and the dash-dot line (-.-.-) shows the path followed by Product C from RM 3 through the various resources. This *resource centric viewpoint* is also the viewpoint of traditional management methods. Cost control is the primary goal of operations management and the traditional view is that resources drain or consume costs. *The way to manage cost is to manage the efficiency of each resource and to make sure that no time is wasted at any resource.* Goldratt (2003) has aptly captured this viewpoint of traditional operations management in the phrase: “A resource standing idle is a waste.” Consistent with this view, most measurements in operations are resource centric (local departmental measures such as efficiencies, utilization, downtime, etc.) and are designed to capture information on what resources were doing every second of the day.

An alternate viewpoint of the same factory floor is to *look at how materials flow. Materials move through the factory, flowing from raw material to finished product. Along the way, they are transformed or worked on by resources. The material thus flows from raw material to one resource to another until it is fully transformed and the finished product leaves the factory or operation.* We call this viewpoint a *flow centric viewpoint*. From a *flow centric viewpoint*, the same factory in [Fig. 8-1](#) would be represented as shown in [Fig. 8-2](#). Since there are three separate materials, there are three separate flows. The transformation of any product, such as Product A (represented by the solid line), from raw material (RM 1) to finished product can be represented by a unique sequence of operations—resource R1 performing operation 010, then resource R3 performing operation 030, etc. We have chosen the vertical direction to depict the time sequence of these steps. In this particular case, the three separate products are produced in very similar fashion; they

follow identical paths. [Figure 8-3](#) shows the resource centric and flow centric views when the three products have very dissimilar routings. These diagrams are referred to as *Product Flow Diagrams* or PFDs.

The essence of the manufacturing operation—the transformation from raw material to a finished product—is reflected in the flow centric view. It is not surprising that the management of production/manufacturing operations should be based on a flow centric view and not a resource centric view. In his article, “Standing on the Shoulders of Giants,” Goldratt (2009) presents the core argument that Henry Ford’s assembly line process and Dr. Taichi Ohno’s Toyota Production System (TPS) originate from a focus on flow.

By a flow centric view, we mean much more than looking at production operations in the format of [Fig. 8-2](#). The primary role of production management is recognized to be to manage flow. Effective management of flow implies that the movement of all material through the factory will be smooth and fast with no stoppages. In any flow, obstacles to the flow result in buildup of material—a traffic jam—and are considered highly undesirable. Resource centric methods are interested in keeping resources busy and consider buildup of material unavoidable.

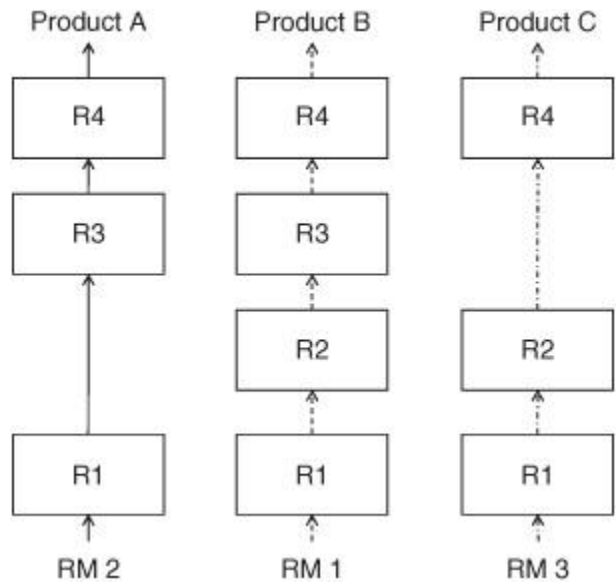


FIGURE 8-2 A flow centric representation of the plant in [Fig. 8-1](#).

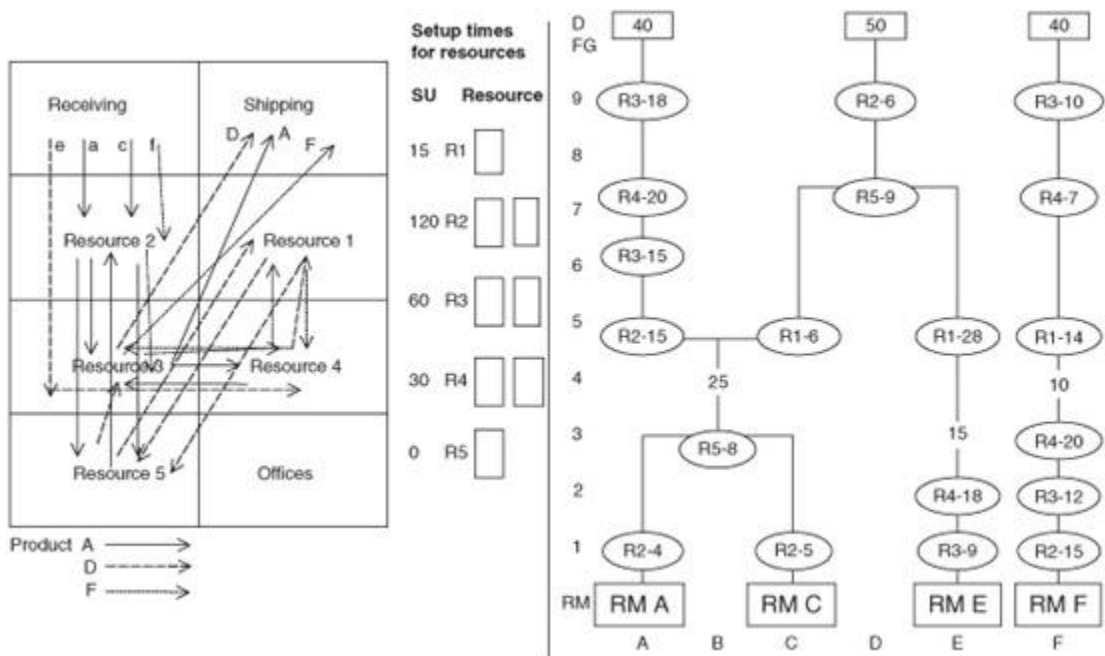


FIGURE 8-3 A comparison between the resource centric representation and the flow centric representation of a plant where the routings for the products are dissimilar. [© E. M.

Goldratt used by permission, all rights reserved. Source: Modified from E. M. Goldratt (2003, 29)]

To understand the key difference between the two views, think of it this way. When walking through the factory you are bound to see idle resources and idle batches of material. Which bothers you more in the pit of your stomach? If the resource standing idle bothers you more, then you are exhibiting a resource centric view. If the batch of material sitting idle bothers you more, then you are exhibiting a flow centric view. What we have learned from Henry Ford and Taichi Ohno is that the flow centric view is the proper point of view for effective management of the system.

Traditional cost accounting-based management methods, unfortunately, are resource centric in nature. Operators typically describe themselves in terms of the resource or resources they operate—press operator, furnace operator, etc. Managers also describe themselves in terms of the resources they control—Press Department, Heat Treat Department, etc. The entire management control system is geared to track the activities of the resources and, in particular, to track, understand, and hence eliminate non-production or idle time on the resource.

Ford and Toyota Production Systems—A New Perspective

In a groundbreaking article in 2009, Goldratt provided a new perspective on the two production methods that have defined this field in the last 100 years—the Henry Ford assembly line system and the Dr. Ohno TPS. Everyone knows that Henry Ford is the father of the modern assembly line mode of production, but most have focused on how this system enables better utilization of resources (material is brought to the worker) and it achieves the dream of balancing capacity. Goldratt took a different point of view.

Henry Ford's real objective was to improve the flow of products through his factory. He was so successful at improving flow that in 1926, the elapsed time between unloading iron ore from the boat to the same iron being loaded on a freight train as a finished automobile was an astonishingly fast 81 hours (Ford, 1928). The magnitude of this achievement is underscored by the fact that eight decades later, no automobile manufacturer can come close to Ford's achievement. Contrary to the traditional belief that one cannot achieve maximum or full output without ensuring that all resources are productive and producing all of the time, Ford's method produced far more output from the factory as a whole. In fact, a focus on flow can result in some resources running out of work occasionally. However, system Throughput is not compromised, but actually enhanced. The success of Henry Ford clearly demonstrated that the resource centric view had led to false assumptions, but this lesson was mostly lost in history from 1926 to the 1970s and the emergence of TPS.

Goldratt concludes that both Henry Ford's assembly line and Taichi Ohno's TPS were systems in which achieving smooth flow through production was a prime objective and that the generalized method they followed can be summarized by the following four principles:

1. Improving flow (or equivalently lead time) is a primary objective of operations.
2. This primary objective should be translated into a practical mechanism that guides the operation when not to produce (prevents overproduction).
3. Local efficiencies must be abolished.
4. A focused process to balance flow must be in place (Goldratt, 2009).

Of particular significance is the second statement. Goldratt points out that the assembly line and the Kanban system of Toyota are essentially systems that tell workstations when not to produce. For instance, in an assembly line if one workstation stops, all others have to stop because the line stops and there is no place to put material if any of the other stations continue to produce. Similarly, in a Kanban system when there are no Kanban cards, work centers stop working. In contrast, in most traditional production operations one of the key arguments for maintaining significant work-in-progress (WIP) queues is to decouple each work center from other work centers and their possible disruptions.

Henry Ford relied on space to limit production while Taichi Ohno developed the Kanban system³ to do the same. Of course, if we are introducing a system that intentionally stops resources from producing, then clearly Principle 3 (abolish local efficiencies) is unavoidable. What is interesting is that both Henry Ford and Taichi Ohno did not simply stop at limiting production, but leveraged these situations into opportunities for improving processes that streamlined and increased the volume of flow. When the built-in mechanisms—space or inventory—create a line stoppage, one has clear visibility about what caused the stoppage and, hence, points to the problem that needs to be solved to better balance the flow. The magnitude of improvements that both Ford and Toyota were able to achieve over their competitors in increased speed and reduced total cost stands as testimony to the effectiveness of their approaches.

In spite of the tremendous success of their methods and the volumes of articles and books written about their methods, the focus on flow did not spread to all parts of the manufacturing industry. In a small country like Japan, given the clear success of Toyota as a business and their

attribution of this success to TPS, one would expect wide adoption of TPS. In fact, less than 20 percent of the manufacturers have implemented TPS and few of these manufacturers have achieved Toyota's level of success. What is the cause for this low level of adoption and success? Certainly, it is not lack of desire or knowledge. Almost every company has attempted to adopt TPS or Lean Production as it is also known. There is an ocean of material on TPS and Lean and Toyota has been very open about its techniques. Goldratt (2009) concludes that the core issues are twofold:

1. The resource centric mindset is still the prevailing viewpoint. This explains why even when TPS is applicable and is adopted, the results are less than what is possible.
2. The specific mechanisms for preventing overproduction—space in the case of Ford's assembly line and inventory in the case of Ohno's TPS—are not applicable to all manufacturing environments.

In his article, Goldratt proposes a different and more universal mechanism for preventing overproduction. He proposes the use of time. To prevent overproduction or producing early, one should not make the material available early. Exactly how we determine the time when material should be released and the additional rules for managing flow are described in the following.

Production Operations and the Five Focusing Steps of TOC

In this section, we discuss the application of the core principles of TOC to production operations. As discussed in other chapters, the 5FS provide the rules for determining how any given operation should be managed. These steps (Goldratt, 1990b, [Chapter 1](#)) are listed below:

Step 1: Identify (or choose) the system constraint.

Step 2: Decide how to exploit the system constraint.

Step 3: Subordinate all other decisions to the above.

If we desire to improve the performance of the system to a level higher than possible with the current constraint, then we must

Step 4: Elevate the system constraint.

This step can change the constraint or the decisions on how to exploit the constraint. Hence, the need for Step 5.

Step 5: If, in Step 4, the constraint is broken, then go back to Step 1. Don't let inertia become the constraint.

Production is only a part of most manufacturing business organizations, that is, it is a subsystem. The true constraint of the business may or may not be in the production subsystem of the organization. If the constraint is chosen to be another subsystem or the market, then the role of production in the five-step process is under Step 3—Subordination. In this case, production should be managed by the rules of *Simplified Drum-Buffer-Rope* (S-DBR) discussed in [Chapter 9](#).

The other possibility is that we choose the constraint to be in the production operation. More specifically, the capacity of a specific work center is chosen to be the constraint. By this choice, the company is making the statement that its business strategy is to make money by finding the best ways to exploit the available capacity at this work center. Clearly identifying the specific resource that will be the constraint (Step 1) and then finding the rules to exploit the capacity of this constrained resource (Step 2) are the key elements of the DBR system for managing production operations.

In this section, we present the DBR system for managing the flow of products in production operations. The scope of decisions that are involved in exploiting the constraint goes far beyond managing the flow of products. For example, the choice of which products to market has a significant impact on the total Throughput potential of the factory. An excellent discussion of this case (referred to in the TOC literature as the PQ example) can be found in *The Haystack Syndrome* (Goldratt, 1990a, [Chapters 11-13](#)) and is also discussed in [Chapter 13](#) of this book. For our purpose, it is assumed that we know what products are being sold and who the customers are. The challenge we are addressing is how best to manage the flow of products so we are able to satisfy this customer demand while keeping inventories and expenses to a minimum.

Characteristics of Production Operations

Every production operation is characterized by the following elements.

There Is a High Degree of Dependency

Dependency in this context means that certain operations or activities in the plant cannot take place until certain other operations or activities are completed. Some examples of dependency in a manufacturing operation are as follows:

- The routing sequence of required operations to manufacture a product is a simple example of manufacturing dependencies. In the typical case, the production process cannot begin until the required materials have been procured; individual operations cannot be performed until the prior operation specified in the routing has been performed; and the assembly

operation cannot begin until all required components have been fabricated or purchased.

- Another obvious example of dependency is the same resource being required to process more than one operation. These operations can be different steps in the routing of the same product (rough milling and finish milling, for example) or steps on different products (rough milling of Product A and rough milling of Product B). The possibility for creating blockage for one product when the resource is occupied with another product is obvious.

Other examples of dependency include:

- Resources cannot be set up until the setup person is finished with another job.
- Work cannot begin until the setup or changeover is complete.
- The first piece of a lot cannot be inspected and approved until the inspection gauges are calibrated.

The number of dependencies in even a small production operation is staggering.

Production Operations Are Subject to a High Degree of Variability

Variability exists in manufacturing operations in the form of both random events and statistical fluctuations. *Random events are those activities that take place at irregular intervals, have no discernable pattern, and by nature are unpredictable.* Examples of random events include:

- A significant customer order is suddenly cancelled.
- A key vendor's plant is crippled by a strike and the critical materials are not readily available.
- Tools, fixtures, gauges, etc., are suddenly unavailable due to unexpected breakage.

Statistical fluctuation or common cause variations in manufacturing environments refer to the fact that *all processes have some degree of inherent variability*. Examples of statistical fluctuations include:

- Receipt of materials from vendors can vary in quantity, quality, or timing from purchase order to purchase order.
- Time to set up a resource varies each time the resource is set up.
- Actual customer orders are different from the forecast.
- Process yields may change from one lot to another.

We will use the term *variability* to describe *both random events and statistical fluctuations*.

The existence of these two phenomena—dependency and variability—combine to make the task of controlling the performance of manufacturing operations very difficult. In fact, the day-to-day role of a shop floor manager is nothing more than attempting to cope with the almost endless stream of disruptions and their impact on a wide range of activities.

At a single step in any process, it is not safe to assume that the effect of statistical fluctuations will average out and the performance of the process will be the average rated performance for that step. One of the dramatic effects of

having both dependencies and fluctuations is that this averaging out does not occur. As discussed in detail in several other works (Goldratt and Cox, 1984; Srikanth and Umble, 1997; Schragenheim and Dettmer, 2001). “Disruptions/fluctuations will not average out for the total system and most individual resources will be forced to perform below their capability” (Srikanth and Umble, 1997, Vol. 1, [Chapter 4](#)).

Resource Capacities Are Unbalanced to Each Other and to the Market Demand

The ideal goal that every operation strives to achieve is that of a balanced capacity plant—every resource has just enough capacity to meet market demand. A major effort of most manufacturing operations is to manage the capacity that is available so that there is no wasted or excess capacity. In spite of this enormous effort, the perfectly balanced plant does not exist in reality. This is due to two factors. The first factor is that capacity comes in finite increments—resources must be purchased in whole units, labor must be hired for one full shift, etc. Thus, if we need 2.67 units of a particular resource we have to end up with 3 units.

The second factor that makes it impossible to have the ideal perfectly balanced plant is the combined effect of dependency and fluctuation. As discussed in the previous section, resources downstream will feel the impact of disruptions in upstream processes in a very biased fashion—they feel the impact of negative variations, but not those of the positive variations (see Srikanth and Umble, 1997, Vol. 1, [Chapter 4](#)). As a result, resources downstream will fall further and further behind, unless they have available capacity to catch up. If the plant were perfectly balanced, there would be no catch up capacity available and the plant would fall further and further behind. Without an

appropriate amount of reserve capacity, the plant will be unable to operate effectively. As the plant falls behind schedule, managers will be forced to increase capacity (through overtime, hiring additional labor, etc.) at the resources that have the most delays. Thus, in the end, managers are forced to run unbalanced plants.

The total available capacity of a resource can be broken down, based on the previous discussion, into three categories: productive capacity, protective capacity, and excess capacity.

Productive capacity is defined as *resource capacity that is required to produce a quantity of product sufficient to satisfy the agreed upon output of the system* (Sullivan et al., 2007)⁴. *Protective capacity* is the *resource capacity needed to protect the Throughput of the system by ensuring that some capacity (above the capacity required to support system Throughput) “is available to catch up when disruptions inevitably occur. Non-constraint resources need protective capacity to rebuild the bank in front of the constraint or capacity constrained resource (CCR) and / or on the shipping dock before Throughput is lost”* (40). *Excess capacity* (22) is defined as *resource capacity that is in excess of what is required to protect Throughput of the system*. Protective and excess are also called idle as most of the time they are not used; protective engages when Murphy strikes to rebuild buffers.

It is a far better strategy to acknowledge that perfectly balanced plants are not attainable and are not even desirable. This means that most real life production operations are unbalanced and many resources will have idle (composed of protective and excess) capacity available. The availability of this idle capacity allows us to design a system under which the operation as a whole will perform at a higher level of reliability (less fluctuation) than individual operations.

Applying the Five Focusing Steps to Production Operations

We are now in a position to design a system that can operate at a very high degree of reliability while producing the highest levels of output possible. Since we do not have a balanced plant, it is clear that at least some resources will have more capacity than needed to meet market demand. In fact, in any dependent chain of resources there will be one resource that has the least capacity relative to demand. If the capacity of this resource is the same or less than the capacity required to meet market demand, then the resource is referred to as a bottleneck. The weakest bottleneck is the constraint of the system.

The rules one must use to get optimal performance from any system are derived under TOC through the application of the 5FS. The resulting approach is referred to as the DBR method of managing production operations. The application of the 5FS would proceed thusly.

Step 1: Identify the System Constraint

In this case, we are dealing with a situation in which the constraint is the available capacity at a resource. The simplest way to identify such a constraint is to compare the load placed on each resource with the total amount of production and setup required at that resource to satisfy the market demand. However, this does not always produce meaningful results due to inaccuracies in data. In several hundred factories with which the author has consulted, this method fails to identify the real bottleneck in an overwhelming number of cases. Detailed procedures for identifying the bottleneck have been developed for each of the different production flows—V, A, T, and I—and are briefly discussed at the end of this chapter. (See Srikanth and Umble, 1997, Vol. 2, [Chapter 4](#) for V-plants, [Chapter 5](#) for A-

plants, and [Chapter 6](#) for T-plants.) The choice of the bottleneck is the pivotal point in the development of the strategy for the entire business and hence this is a decision that must be made by the business as a whole and is not just a production/manufacturing decision.

Step 2: Decide How to Exploit the System Constraint

The constraint in the environment we are discussing is the available capacity at a specific resource. Exploitation of that resource means that we should maximize performance with respect to the global operational metrics of Throughput, Inventory, and Operating Expense. More specifically, the goal is to maximize Throughput, while efficiently managing Inventory and Operating Expense. How can we maximize the Throughput of the production operation with a specific capacity constraint or bottleneck? To answer this question, we can look at ways in which capacity is currently wasted.

By definition, the load placed by current market demand on a bottleneck is greater than or equal to the available capacity at this resource. If the resource spends any time doing something other than what is required for current market demand, then Throughput will be negatively impacted and we will not have properly exploited the available capacity. It is, therefore, critical that every item produced at the constraint be a product that is required to fulfill short-term market demand. Another way in which capacity at the constraint can be wasted is for the resource to suffer a breakdown and then a significant time to elapse between the resource breakdown and its being fully operational. Excessive setup times, time lost during shift changes or lunch breaks, etc., are all ways in which capacity at the bottleneck is wasted and represents the opposite of exploitation. Policies, such as overlapping shifts and staggering breaks should be put in place to eliminate these

forms of wasted capacity. Capacity is also wasted when the bottleneck works on products that are not needed to satisfy current market demand. While this may appear to be so obvious as to constitute a triviality, the reality in most operations is very different. Motivated by local optima considerations, often bottleneck resources end up working precisely in this wasteful way—because either no other work is available or batch sizes in use are excessive. One of the prime considerations in designing the rules that will allow proper exploitation of the bottleneck is to make sure that the bottleneck does not run out of work and that the planned work (as well as actual material available on the factory floor) consists only of the product required to meet very near term demand. The procedures for doing this are discussed in the section on the drum.

Due to the existence of dependencies in manufacturing operations, the performance of any resource is influenced by the performance of other resources. In the simple flow shown in [Fig. 8-2](#), resource R4 cannot continue to work if resource R2 is down for an extended period of time. If resource R4 is a non-bottleneck, then the forced downtime at R4 is not a serious issue. If, however, R4 is the bottleneck in this production flow, then the downtime at R4 is unacceptable as system Throughput is reduced. To ensure that resource R4 can continue to work even when upstream flows experience disruptions, we must maintain a buffer at R4 (enough work to cover the time the resource is down). Since the objective of this buffer is to protect R4 from upstream disruptions, the size of this buffer is a function of the magnitude and frequency of these disruptions. While determining the “optimum” buffer size is quite complex, the two limits are obvious—the buffer should not be so small that the bottleneck is frequently at risk of running out of work and the buffer should not be so big that the total lead time for the flow is excessive. In the section on buffers, we

discuss the procedure for how to set the size of these buffers.

Step 3: Subordinate Everything Else to the above Decision

Once the bottleneck or capacity constraint has been identified, policies to ensure full productive utilization is put in place, the resource properly buffered, and the planned flow through this resource is identified, then Step 2 is complete. The next step, *Subordinate*, is to make sure that all other resources focus on performing tasks in such a way that the planned flow through the constraint is supported.

All activities from the release of material to how they are processed upstream and downstream of the bottleneck should be done in a manner that best supports the decisions made in Step 2. It is important to realize that while the discussion on the constraint is powerful and interesting, the task of execution falls mostly on Step 3 and the management of non-constraints. This is a simple consequence of the fact that most resources (95 to 100 percent) are non-bottleneck resources and to control execution means controlling what is happening at these resources.

The subordination required by Step 3 is made challenging because the mentality fostered by traditional cost world management is not consistent with subordination. This is the point at which the third principle of flow management (efficiencies must be abolished) needs to be implemented. The case in [Fig. 8-4](#) illustrates this point. Resource R2 is the constraint and can process 100 units per day. Resource R1 is a non-constraint and can process 120 units per day. Subordination requires that R1 only process 100 units per day, but traditional mindset would encourage R1 to work to its full potential and thus produce in excess of 100 units per day. In almost every implementation of DBR that the author

has done, the task of subordinating (or holding back production) non-constrained resources has been the most challenging and difficult task.

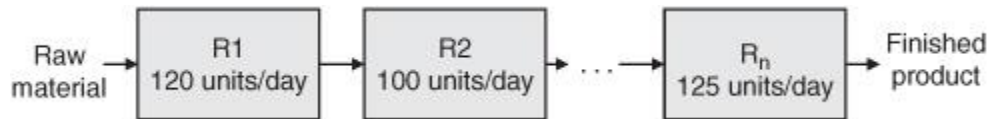


FIGURE 8-4 PFD for a one-product flow line indicating the production capacity of the different resources.

An alternate way to look at Step 3 is as follows. Steps 1 and 2 have established the total flow that must be achieved—product mix, volumes, etc. In accordance with the section on managing flow, we must now implement the four principles of flow. In particular, recognizing that improving flow is the primary objective, we must establish how to implement Principle 2—a mechanism to prevent overproduction. This overproduction (Sugimori et al., 1977) is the first and most important waste that is explicitly identified in TPS as well as in JIT, Lean, and other offshoots of TPS.

The process by which subordination is enforced in the DBR system is the rope and is discussed in a later section.

Steps 4 and 5

At the completion of Step 3 (subordination) we have a system that is operating at full potential—we are getting the maximum Throughput by what has been done at the constraint and waste is minimized by subordination at all other resources. In order to improve the performance of the system further yet, we must raise the performance of the constraint itself. However, when the performance capability of the current constraint is elevated, it may no longer remain the constraint. Its new potential may be larger than

the capability of another resource in the system. Steps 4 and 5 are designed to deal with this possibility. Since our focus is managing a plant that has a current constraint, we will not discuss ramifications of Steps 4 and 5. Rather, we proceed to the implementation of Steps 2 and 3 through the DBR mechanisms.

The DBR System

We now discuss the specific procedures and methods that make up the DBR system for planning the flow of product through manufacturing operations. BM is the execution control portion of the DBR system. The objective of the DBR system, as for any planning and control system, is to meet Throughput expectations while efficiently managing Inventory and Operating Expense.⁵ The essence of the DBR approach is captured in [Fig. 8-5](#).

The Drum

The drum considers the constraints in the system and firm customer commitments, in setting the pace for the entire system. The process of setting the drum begins by identifying the work that needs to be done at the constraint by the total output required. In the case of companies that are make-to-order (MTO), this is the work required to be performed by the CCR to meet all customer requirements that fall in a given time period (for example, all orders with customer due dates in the next 30 days). In the case of make-to-availability (MTA)⁶ companies, the output requirement is the total finished products required to fill the stock buffers. Once we have a list of what must be produced at the constraint, it is then simply a matter of determining the sequence of production (which product first, which

product next, and so on) and the production batch size (how much will be produced once we start a specific product). Factors that should be considered in deciding the production sequence and the size of the process batch as well as detailed examples are found in Srikanth and Umble (1997, Vol. 1, [Chapters 7](#) and [8](#)) and Schragenheim and Dettmer (2001).

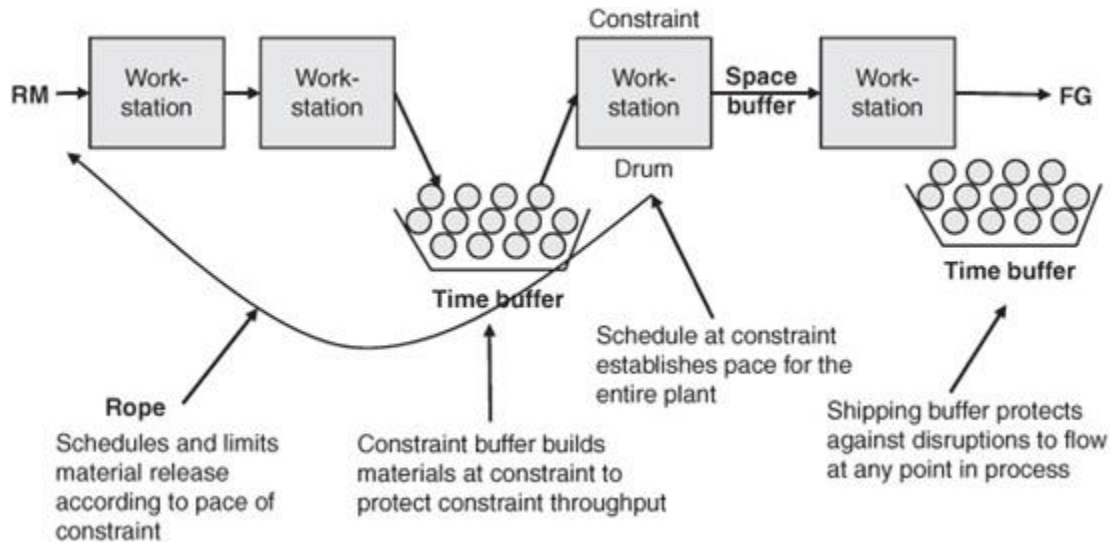


FIGURE 8-5 Illustration of the basic DBR system.

The Buffer

In a world free from disruptions, such as resource breakdowns, process yields, etc., *the production lead time—the time that we allow for the raw material to be transformed to a finished part or product—can be simply equal to the sum of the process times and setup times at each step of the routing for that product.* In the real world where there are many forms of disruption, the use of a planned production lead time equal to the sum of processing and setup times would be considered foolish and rightfully so. Any disruption such as a resource breakdown

would make it impossible to produce the product on time. The actual production lead time will always be larger than the sum of process times and setup times. Since disruptions are unavoidable, planned lead times will have to be larger than the sum of process and setup times. This is true if we are to have any chance of making the *actual* production lead time equal the *planned* production lead time.

Whenever there is a task that is subject to variability, it is clear that the actual time the task is executed—started or finished—is going to be different from any plan that does not allow some degree of padding in the form of safety time. This is essentially the concept of the time buffer.⁷ What makes the application of the time buffer concept unique and powerful is the explicit recognition that the goal of a DBR planning system is not to make each task to be on time to a planned schedule, but to make the actual flow through the system sufficiently reliable to satisfy market demand. In other words, the objective is not to protect the ability of each task to be on time (to a plan) but only to make sure that the entire system is on time. This recognition allows us to provide a significantly higher degree of reliability in a DBR plan than one that tries to ensure protection for each step in the process (as in a push system or Kanban pull system). In addition, this higher degree of reliability can be accomplished at a significantly lower production lead time.

Specifically, a *time buffer* is defined as follows. A *time buffer* represents the additional lead time allowed, beyond the required setup and process times, for materials to flow between two specified points in the product flow. Two points⁸ commonly used in this context are material release (gating operations) and receipt of a finished product at a warehouse (MTA) or at shipping (MTO). The objective of these time buffers is to protect the system Throughput from the internal disruptions that are inherent in any process. The

relationship between production lead time and process times can be expressed by the following relationship.

Production lead time = Sum of process times and setup times + Time buffers

The concept of time buffers is almost self-evident. Determining the proper size of a time buffer on the other hand appears to be a complex task. Since the objective of the time buffer is to protect the flow through the system from disruptions, it might appear that a detailed knowledge of these disruptions—the statistical distribution curves at each step in the flow—would help (or even be a necessary element) in calculating the size of the time buffer. While this may appear to provide a rigorous methodology, it is practically useless because the required information is not available. In the practical application of DBR, we take a more pragmatic approach to establishing the size of the time buffer. Every production operation currently uses a time buffer, whether or not it is explicitly understood. By this, we mean that the production lead times used—informally or in a computerized ERP system—is many times larger than the process and setup times. All of this additional time is a time buffer. We also know that most often this currently used time buffer is much too large. This is because buffers are used to protect each step in the flow and not just the system flow as a whole and, more importantly, larger time buffers make it possible to minimize the situations in which resources simply run out of work. Since the traditional view is that a resource standing idle is a waste, lead times must be large enough to minimize idle time at each resource. In effect, one can view the current lead time as giving us an upper limit—the point where current lead time provides too large of a time buffer.

If current lead times establish one extreme for the time buffers, another extreme—a time buffer that is too small—is provided when the production lead time is close to the sum

of process and setup times. In fact, in almost all production operations a production lead time that is even just three times the process time would be considered unrealistically aggressive. At each extreme end, the time buffers are actually ineffective in providing protection and promoting smooth reliable product flow. When the time buffer is too small, the cumulative disruptions that every batch of product is subject to quickly overwhelm and consume the available buffer. When the time buffer is too large, the shop floor is clogged with too much material, making it difficult to manage the flow. Each operation will have plenty of work from which to choose, and the chance that they will all be coordinated to choose the right work to promote a smooth organized flow is slim. The results are piles of inventory everywhere, long lead times, poor due date performance, and chaos on the shop floor. Between the two extremes is a range of options. Based on vast experience, we believe that [Fig. 8-6](#) captures the essence of the effectiveness of time buffers as we increase buffer size from very small to very large. The key observation is that the curve in the region where the time buffer has a high degree of effectiveness is relatively flat. This means that there is no real benefit to complex calculations that yield precise buffer values. Being in the right ballpark is sufficient. Again, based on vast experience, a good value for the time buffer in most production environments is one-half of their current production lead time.

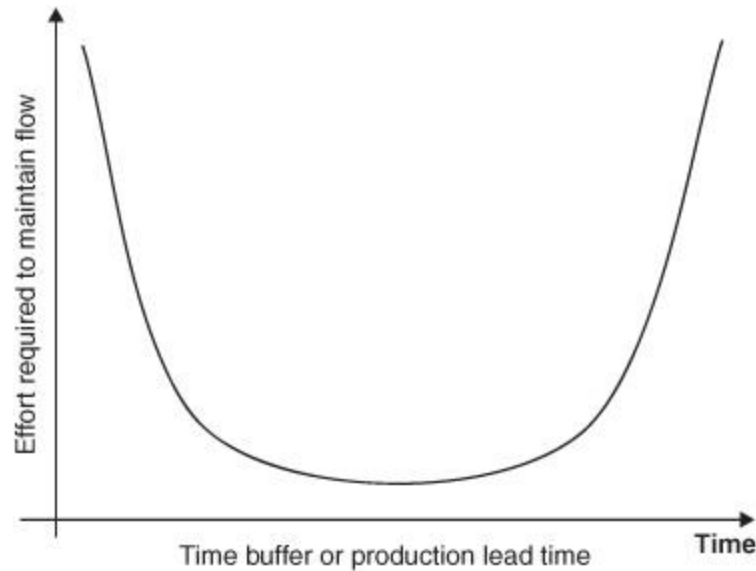


FIGURE 8-6 Graphical representation of the effort required to maintain a smooth flow as the buffer is increased.

The time buffer established here becomes the time element that will be used to implement the second principle of flow management (prevent overproduction). If we want to prevent production ahead of time, then we should *not* make the material available ahead of time. The time buffer provides the amount of the time to be used in Principle 2—a mechanism to prevent overproduction—and the rope mechanism discussed later will enable us to enforce this principle.

Whether a production operation has a true bottleneck or not, as long as there are disruptions the need for a time buffer exists. The only way to ensure that the flow at the end of the system meets promised due dates is to provide protection from disruptions using time buffers. When there is a bottleneck in the system, and this is the case when we are dealing with the full DBR system, there is a need for an additional level of protection. Any time lost at a bottleneck, by the very definition of a bottleneck, will be Throughput lost for the entire system because this lost time cannot be recovered. Hence, if one hour is lost at the bottleneck, then

effectively the total system will be down for one hour and we lose the Throughput that would have been generated during this time. The downtime at a bottleneck can arise from problems at the bottleneck itself (downtime, setups, etc.) or from the same problems upstream from the bottleneck. The bottleneck can be decoupled from the disruptions upstream if we can ensure that there is always material ahead of the bottleneck. The amount of material that is sufficient to provide adequate protection depends on the nature and distribution of upstream disruptions. Note that the constraint buffer at the bottleneck is not created by adding more time into the previously established time buffer. Since the bottleneck is the true constraint to flow, material naturally accumulates at this operation/resource. All other resources have protective capacity and should be able to keep the products flowing. However, when disruptions upstream are of such a nature as to prevent the accumulation of material at the bottleneck, they threaten to create downtime at the bottleneck. This must be avoided and can be done during execution control by monitoring the amount of work at the bottleneck and taking corrective action whenever the work queue at the bottleneck is dangerously low.

It is instructive to point out here that there are other types of buffers used in the overall management of flow in a supply chain. In addition to time buffers, three other types of buffers exist: capacity buffers, stock buffers, and space buffers, with respect to production planning and control systems. The *capacity buffer* is defined as *the protective capacity at both constraint and non-constraint resources that allow these resources to catch up when Murphy strikes*. *Stock buffers* are defined as a *“quantity of physical inventory held in the system to protect the system’s throughput. Perspective: Stock buffers should not be confused with time buffers such as the constraint or*

shipping buffers” (Sullivan et al., 2007, 43).⁹ Stock buffers may be used for raw materials, WIP items (for example, at major divergent points in a V- or T-plant), and finished goods items to reduce lead time or protect against product variety. The *TOCICO Dictionary* defines *space buffer* as “*Physical space immediately after the constraint that can accommodate output from the constraint when there is a stoppage downstream that would otherwise force the constraint to stop working*” (Sullivan et al., 2007, 41).⁹ The idea is to keep the space buffer empty in the same manner as one tries to keep the constraint and shipping buffers full. BM should be used on each of these types of buffers to ensure effective operation of the constraint and high due date performance. They should also be monitored to ensure that they are not too large. Time buffers impact lead time, while stock buffers impact inventory investment.

The Rope

The final component of the DBR system is the *rope*—*a mechanism that is used to control the flow through the system by controlling the flow at a small number of control points*. The drum has created a master schedule that is consistent with the constraints of the system and is best able to satisfy customer demand. The time buffers provide the safety or insurance that the flow to the market will be reliable in spite of the impact of disruptions. The last link is to communicate effectively to the rest of the operation the actions that are necessary to support the drum and to ensure effective control of these actions.

The basic challenge is to ensure that all work centers perform the right tasks in the right sequence and at the right time. With computers becoming ubiquitous in manufacturing, it is very tempting to accomplish this objective by providing each work center with detailed

schedules that are constantly updated (hopefully in real time). DBR takes a counterintuitive but far simpler approach to accomplishing this goal. The simplest and most effective way to make sure that a work center does the right job is to have only the material for the right job available. Eliminate unnecessary WIP and you eliminate opportunities for working on the wrong stuff. With this approach, the emphasis of control is shifted to strictly limiting material available at a work center to what is immediately needed. In production operations, the availability of material in the shop is controlled by the actions at the material release points—the points where raw material is released to fabrication, finished parts are released to assembly, purchased parts are released to assembly, etc.

To implement the rope, the material release points are provided with a detailed schedule that lists what materials need to be released, in what time frame, and in what sequence. If this task is managed properly, then access to unnecessary work is denied to most work centers, thereby forcing them to work on the right products. Most of the work centers that are non-constraints will simply process material when it becomes available. When a work center (a non-CCR) finds itself with more than one batch of material, what are the rules for determining the priority sequence? The real question to ask here is whether sequence really matters. In the majority of production operations, the processing time for a batch of products at any single work center is a very small fraction of the total production lead time or the total time buffer. This being the case, the difference between working on one batch before another is insignificant. It should be remembered that we are talking about very few cases where multiple batches will be available to choose from and even here, the number of batches is small. Thus, a simple rule will suffice to ensure major distortions are avoided. The priority rule can be a simple “first-in, first-out” or FIFO rule.

In simple linear flows, simply controlling the *release of material* will be sufficient to control the execution through the whole system. The basic principle we have followed is that we can make sure that a work center cannot work on the wrong product if material is not available. In other words, the mere fact that material is available is sufficient information to give the green light to that work center for processing. In complex flows, this basic fact is not always true. For example, at *divergence points* (see V-plant discussion later in this chapter), *the same incoming material can be processed into different outgoing materials*. It is obvious that when such a work center can be activated by material availability, we have to specify what the output products should be and how much of each product we want. While the timing of the jobs is controlled by the availability of material, workers at each divergence point (a control point) need to be provided with a detailed list of what and how much of each product to produce, as well as the priority sequence for the products.

Similar to divergence points, *assembly or convergent points* may also need to be controlled. Purchased parts may be obtained in quantities larger than required for specific orders; fabrication may also have combined different orders to reduce setups at CCRs; and in T-plants (see the discussion later in the chapter), *the same basic component parts can be assembled in different combinations to create different end items*. The assembly departments should operate to a priority list that specifies what units should be assembled in what quantities and in what sequence.

The question often asked is, “What about at the *CCR or bottleneck?*” Is a detailed schedule specifying sequence and quantity of production necessary and should this be carefully monitored and controlled? If the CCR or bottleneck has sequence-dependent setups, then the setup time depends on what is currently on the resource and what product is next. A simple case is an operation that applies

color. Going from a light color to a dark color requires minimal cleanup, but going from black to white will require extensive and thorough cleaning. Therefore, it is important to produce to a defined sequence and this list will have to be provided to the CCR. If this is not the case, and the process times are a small fraction of the total production lead time, then sequence even at the CCR is not that critical and no additional step beyond controlling material release is necessary. [Figure 8-7](#) shows the schedule control points where sequence and a time frame for actions are important in controlling the flow through a factory.

Finally, there is the *shipping or completion point* of the batch. It is the most important schedule control point in that every batch has to meet the date when it is scheduled to be completed. Failure of a batch to meet this date or even the anticipation that a batch might fail to meet this date will trigger corrective actions as described in the section on BM.

Managing Flow with DBR—An Example

We illustrate the DBR system with a simple example¹⁰ in this section. The plant represented by the PFD in [Fig. 8-8](#) shows a relatively simple plant with five different types of resources—each pattern in the diagram is a different type of resource (labeled R1, R2, R3, R4, R5). The number in each box in the flow diagram is the time to process one unit at that step. For example, the first step (A-1 on the bottom left of the grid used in [Fig. 8-8](#)) is performed by the R2 resource and takes 4 min for each unit. Similarly, the step corresponding to the grid point B-3 is an assembly operation performed by the R5 resource and takes 8 min per unit to assemble. To assemble a unit at B3, a unit must be completed at both A1 and C1. That assembly can then be used at the A-5 operation to make Product A or by C-5 to make Product D. The number of units of a specific resource

type is indicated on the left-hand side of [Fig. 8-8](#) and shows that there is only one resource of type R1. We also see that there are two resources of type R2. Similarly, we have two resources of types R3 and R4 and one resource of type R5. The setup time for each resource is indicated next to the resource—R1-type resource has a setup time of 15 min, R2-type resources have a setup time of 120 min, and so on. The number just below a node in the flow centric represents the number of units that are available for the resource (the WIP) at this point in time—there are 15 units available for operation E-5 performed by resource R1, etc. The demand for the three finished products is indicated at the top of the diagram and is the weekly demand for the product. For the current week, the demand for each product is as follows: 40 units of Product A, 50 units for Product D, and 40 units for Product F.

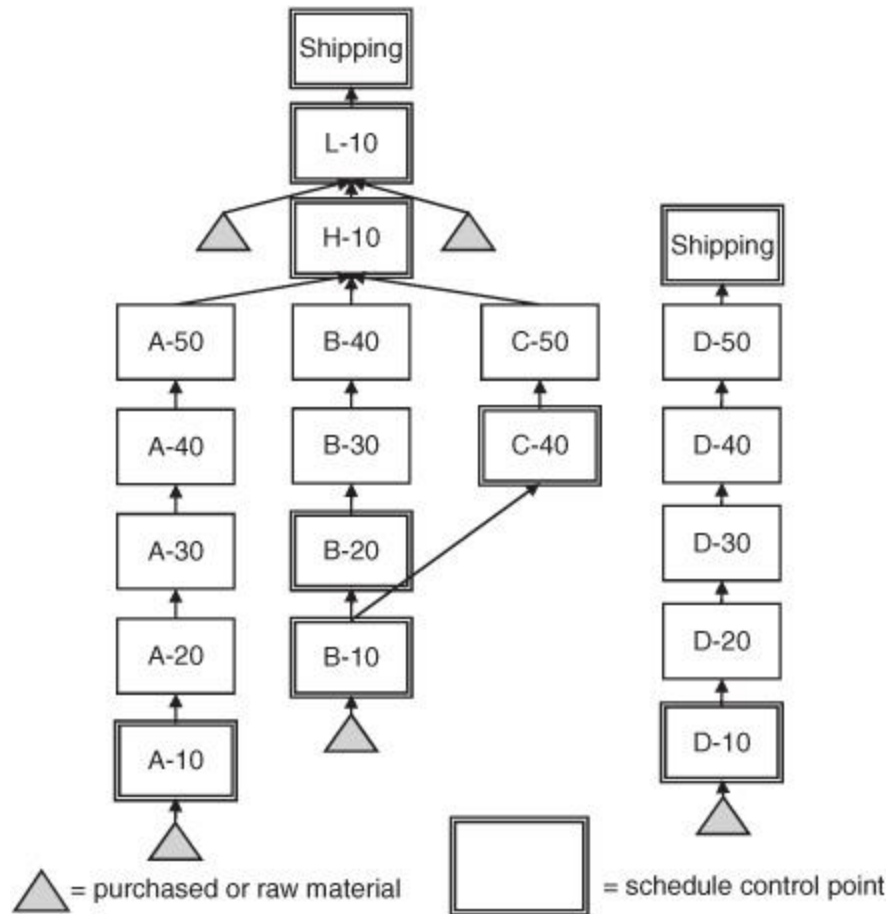


FIGURE 8-7 Schedule control points in a plant with assembly and divergence.

Based on this product structure, the demand, the material in process, and the process information for each product, we can compute the load for each type of resource. For this, we calculate the number of units that need to be processed at a given step and then multiply that by the time to process a unit. For example, at Step A-1, which feeds both Product A and Product D, the total number of units to be processed is 65 (40 units for A + 50 units for D - 25 units of WIP at B-3). The total time required of the R2 resource for this production is (Number of units to be produced) × (Time to produce one unit) or $65 \times 4 \text{ min} = 260 \text{ min}$. In a similar manner, we compute the capacity required at each step that requires an R2 resource; namely, Steps C-1, F-1, and A-

rights reserved. Source: Modified from E. M. Goldratt (2003, 29))

From these calculations, it is clear that this operation has a CCR in the R1 resource. Of the available time, 94 percent is required just to process the units required for the week. The remaining 6 percent is available for setups, maintenance, etc. Any consumption of time (not producing product) that exceeds 6 percent will result in missed shipments. In fact, based on the information that the capacity required for processing the needed parts is 2260 at the R1 resource, we can see that only 140 min (2400 - 2260) are available for doing setups. Each setup requires 15 min. Therefore, we can only incur nine setups. Since there are three distinct steps where we need the R1 resource (C-5, E-5, and F-5), we conclude that we can do up to three setups for each step. To be on the safe side and allow for some fluctuations, we can choose to do two setups at each step. Effectively we will run two batches of 25 units at C-5, two batches of 25 units at E-5, and two batches of 20 units at F-5. [Table 8-2](#) shows a schedule for the R1 resource that is constructed on this premise.

Resource Type	Number of Units Available	Available Capacity per Week (min)	Required Processing Capacity per Week (min)	Percentage Load Required/ Available (%)
R1	1	2400	2260	94
R2	2	4800	1635	34
R3	2	4800	2695	56
R4	2	4800	2310	48
R5	1	2400	970	40

(© E. M. Goldratt used by permission, all rights reserved.
Source: Example based on E. M. Goldratt (2003, 27-36))

TABLE 8-1 Capacity Available and Required and Percentage Load to Satisfy Weekly Demand

Schedule for Completion of Product A	
Quantity	Completion Time (Hours from Zero)
10	16
10	24
10	32
10	40

(© E. M. Goldratt used by permission, all rights reserved.
Source: Modified from E. M. Goldratt (2003, 114))

TABLE 8-2 Schedules for Constraint (R1 Resource) and Market for Product A

Resource

Schedule for R1 Resource		
Task	Quantity	Expected Start (Hours from Zero)
E-5	25	0
C-5	25	12
F-5	20	15
E-5	25	20
C-5	25	32
F-5	20	34

In this example, Product A does not require any time at the R1 resource. *The market is the constraint* for Product A.¹¹ How do we manage the flow of this product? The simplest thing is to produce Product A using the customer orders. However, a single order of 40 units moving through the operation is not an example of smooth flow. To overcome the effects of large and lumpy flow in this case, we will divide the order into four batches of 10 units each and process them to be completed by the end of the week. The schedule for the R1 resource and the schedule of

completions for Product A together represent the drums for this plant.

The next step is to establish the size of the time buffers. For this simple model, we select a constraint buffer of 24 hours (3 days, in this case). In real manufacturing plants, the production lead time currently being used provides the starting point. As indicated earlier, the first choice of the time buffer is to reduce this by 50 percent. For our example, we do not have this reference point. The choice of 24 hours reflects the need for a time buffer that is approximately 20 times the processing time for a unit (i.e., the process time is around 5 percent of the total production lead time). In addition, since all products have comparable routings, the time buffer is chosen to be the same for each. This means that raw material must be released 24 hours (3 days) before the expected completion by the constraint. [Table 8-3](#) provides the rope or the release dates for the various raw materials into the process.

ROPE—Material Release Schedule		
Material	Quantity	Scheduled Release (Hours from Zero to Nearest Half Day—4 Hours)
E	10	0
A	20	0
C	20	0
F	10	0
E	25	8
A	10	8
C	10	8
A	10	16
C	10	16
A	25	20
C	25	20
F	20	20

Schedule for Step A5 (Uses Same Material as C5)	
Quantity	Start Time (Hours from Zero)
20	0
10	8
10	16

(© E. M. Goldratt used by permission, all rights reserved.
Source: Modified from E. M. Goldratt (2003, 114))

TABLE 8-3 The Rope (Material Release Schedule) for Sample Plant

Product	Quantity	Scheduled Completion (8 Hours After Drum/ Control Step)
A	20	14
A	10	18
A	10	28
D	25	23
D	25	42
F	20	28
F	20	47

(© E. M. Goldratt used by permission, all rights reserved.
Source: Modified from E. M. Goldratt (2003, 114–117)).

TABLE 8-4 Expected Completion Times Based on the DBR Schedule

In this example, material release and the divergent point represented by operations A5 and C5 are the only schedule control points and no other information is needed for planning purposes. [Table 8-2](#) (the drum), the choice of 24 hours as the time buffer, and [Table 8-3](#) (the rope) provide the DBR system for this case.

An important question to answer before we commit to execution of the above DBR plan is “Does this plan help us complete the products in such a way as to meet customer expectations?” How can we identify when orders are planned to be completed. For this we have to extrapolate from the expected completion of products at the constraint (—Resource R1 as detailed in [Table 8-3](#)) by adding a reasonable estimate of time for the completion of the remaining steps (from R1 through assembly). In this simple case, we have chosen 8 hours (about one-third of the total planning lead time) for this estimate. This works in this simple case due to the relatively simple nature of the flow (minimal resource or material contention). [Table 8-4](#) shows the time when different batches of Products A, D, and F are expected to be completed and be available for shipment. If we want to commit shipping times that can be met with very high confidence, then we should commit to hours 42 and 47 (which, in a 5-day, 8 hours per day workweek, corresponds to the Monday of week 2). In real-life situations, it is common practice to choose a slightly more conservative estimate and use one-half of the planning lead time. This means that the estimate of when a batch or an order can be completed (and hence available to ship) is equal to the completion of the last needed batch at a constraint plus one-half of the production lead time. In-transit lead time has to be added to this shipping date to determine when the order will be at the customer site.

Managing Flow—Controlling Execution and Buffer Management

The Need for Control and the Need for Corrective Actions

Using the DBR system described previously creates a plan that maximizes the system Throughput, by ensuring full utilization of the constraint while focusing on real customer demand. The plan is robust and protected from disruptions using time buffers and minimizes investment in Inventory by restricting inflow of material through the rope mechanism. This does not mean that the execution of the plan on the shop floor is automatic and that the execution does not have to be monitored carefully. It is true that in creating the time buffers, we have allowed for a certain level of disruption to the flow of a batch of material through the system. As long as the deviations actually being experienced by the batch are less than what was allowed for, we do not have a problem. However, when the actual deviation begins to exceed the allowable disruption, the ability of the batch to reach customers on time will be in jeopardy.

In these cases, not all is lost. In most manufacturing operations, there is opportunity for corrective action to be taken. The objective of these actions is to “make up” some of the time lost by the batch due to larger than anticipated disruptions. These actions include:

- Expediting the batch by moving it to the front of the work queue at each resource
- Working overtime at a resource to process this batch
- Processing the batch on more than one identical resource (batch splitting)
- Overlapping processing (carrying completed materials from one work center to the next to allow both work centers to work simultaneously)
- Alternate routings

The use of time buffers minimizes the need for corrective actions, but it does not eliminate them. What is needed to make the DBR system deliver exceptional results in practice is a mechanism that can identify the cases where corrective action is necessary and help monitor the effectiveness of the corrective actions so that every batch can be finished on time.

Understanding Buffers: The Buffer as the Source of Information for Controlling Execution

In order to identify when a production batch is experiencing larger than “normal” disruptions, we need to go no further than understanding the time buffer in a bit more depth. When a batch of material is released one production lead time before its due date, what do we expect to happen in reality? Let us understand this by studying a sample of 100 identical batches with a production lead time of 40 days. The majority of batches experience disruptions that are within a normal but wide range. Most of the batches (about 90 percent) will reach their destination on or ahead of plan—less than or equal to 40 days. For example, some of the batches will experience far fewer than normal disruptions and these could be completed in, say, just 10 days, a time that is much shorter than the planned production lead time. Similarly, there will be a small number of batches that will experience much more than their fair share of disruptions. In the absence of corrective action, these batches will finish well past their due dates based on a 40-day planned lead time—the batches will be late. The distribution curve for the sample of 100 batches is illustrated in [Fig. 8-9](#).

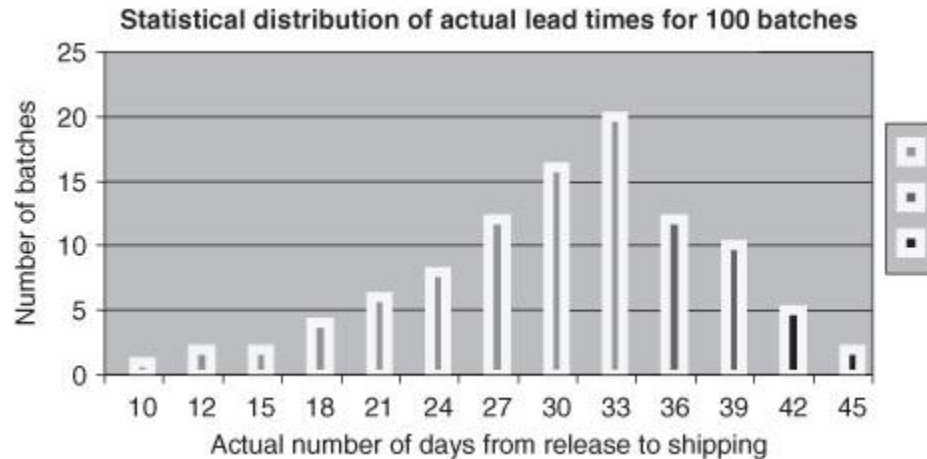


FIGURE 8-9 Graph showing the number of batches with actual lead times ranging from 10 days to 45 days where the planning lead time was 40 days with review at 35 days.

If the only point where we can identify that a batch is experiencing large disruptions is at the end of the product flow, we will have no opportunity for corrective action. We need to know that a batch in production is in trouble while there is still enough time to do something about it. What is the minimum time that will leave us enough time for corrective actions that, in the majority of cases, can help bring the batch back on track? To better understand this, let us consider a batch with a planned production lead time of 40 days that is released today. (Today is Day 1 and Due Date is Day 40.) If we simply let the normal shop floor mechanisms take place without any intervention or without even any monitoring, we expect this order to reach completion sometime between Day 10 (no major problems encountered) and Day 45 (many major problems encountered). Suppose we choose to monitor this order after 35 days have elapsed. From the statistical distribution curve shown in [Fig. 8-9](#), approximately 70 percent of the time the order will already have been completed and the monitoring is a non-issue. However, in the remaining 30 percent of the time the monitoring will reveal the extent of

the disruptions suffered; hence, the urgency of taking corrective action. In many of these cases (approximately 20 percent), the batch will be almost near completion and no action is necessary. In a small number of cases (10 percent), the batch is far behind in its progression through the shop and corrective action will be necessary. We can then initiate these corrective actions and bring this batch back on track.

The general rule that emerges from this example is the following. In trying to determine whether intervention is required, we are comparing two time periods. The first time period, *time available*, is the amount of time that is actually available to finish the batch on time. This is the time from Today/Now to the Date/Time when the batch is due. The second time period, *planning or standard production lead time*, is the amount of time that is required to complete the batch. As the ratio of available time to planning production lead time becomes smaller (this will happen naturally with the passage of time), the degree of certainty that the batch will finish on time will diminish. We refer to the *ratio of available time to standard production lead time, expressed as a percentage*, as the *buffer status of the batch*.

Buffer Status (%) = (Available Time) / (Standard Production Lead Time) × 100%

[Figure 8-10](#) shows the buffer status in the form most frequently used, which is by assigning each work order a color based on the buffer status. If the time remaining for an open work order is less than one-third of its standard production lead time, then the buffer status is less than 33 percent. (If the batch was released on time, then we have less than one-third of the standard lead time available to complete the batch on time.) Such a batch should be flagged. Production personnel will have to investigate where the batch is currently located and determine whether corrective action is necessary. The rule that every batch whose buffer status is smaller than 33 percent should be

flagged for investigation is an empirical rule based on experience. If the point at which a warning signal is issued is too lax (buffer status of, say, 50 percent), then we are likely to receive too many warning signals, creating unnecessary work. Conversely, if the point at which the warning signal is issued is too tight, then very few warnings are issued and, more seriously, there may not be sufficient time to react. Due to the fact that the actual touch times (process + setup) are typically 10 percent or less of the actual production lead time, if we know of serious problems 10 days before the batch is due (for a batch with a standard or planned lead time of 30 days), we should be able to expedite the product to finish on time. In the next section, we discuss how this works in practice, in a production operation that has hundreds of batches in process at any given time.

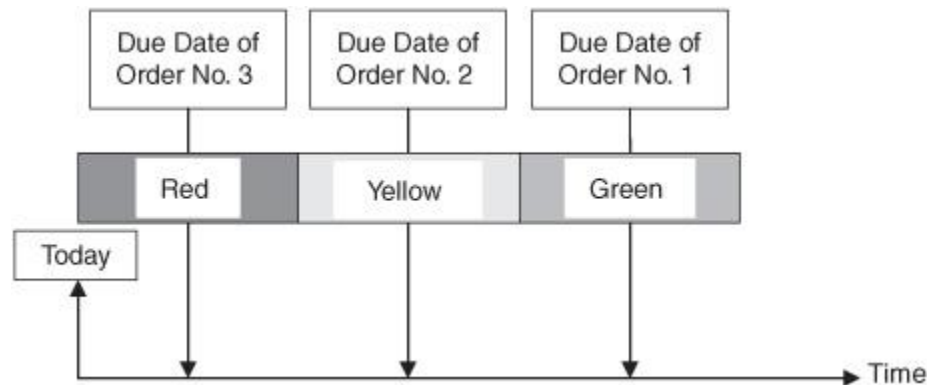


FIGURE 8-10 Designation of buffer status by a color. Comparison of time remaining (Due Date of Order to Today) to the planned buffer time to assign color to a work order. Status and Action: Red—Time remaining less than one-third of the buffer. Expedite. Yellow—Time remaining between one-third and two-thirds of the buffer. Monitor and plan. Green—Time remaining greater than two-thirds of the buffer. Do nothing.)

Buffer Management—The Process

In accordance with the previous discussion, each open work order or production batch will have a buffer status that can be calculated. Note that the buffer status does not depend on where in the process a specific work order/batch is located. Based on the buffer status, the work orders are color-coded into three different categories.

Green Work Orders: *A work order is assigned the color green when the buffer status is greater than 67 percent.* For a green work order, there is plenty of time still available to complete the order. No matter where in the production process this order happens to be, there is no cause for concern and it is reasonable to expect that the order will in fact be finished on time.

Yellow Work Orders: *A work order is assigned the color yellow when the buffer status is between 33 and 67 percent.* For a yellow work order, disruptions have eaten into the normal flow and there is a risk that additional disruptions might make these orders late. However, for now there is no need for intervention.

Red Work Orders: *A work order is assigned the color red when the buffer status is less than 33 percent.* The time left for finishing this order on time is small (relative to what we would like to have as expressed by the standard lead time). It makes sense to see where in the process this order is located. If the order is near completion, no intervention may be necessary. If the order is still in the early stages of processing (or even waiting for material release), intervention is required to mitigate the risk of a late order.

Each work order thus has a color code assigned to it based on the buffer status at that point in time. As time evolves, the buffer status may change. At the beginning of the shift, production managers should construct the list of work orders that are red at that control point. Each of these orders should be investigated to determine if corrective action is called for. Then responsibility for the corrective action should be assigned. The next day, actions should be reviewed to make sure they were done and the new list of red orders should be investigated. In fact, the primary activity of the daily production meeting is the BM process.

The assignment of colors gives us an opportunity to refine the priority system inherent in the DBR process. The simple FIFO rule can be modified as follows: Red orders first, then yellow orders, and then green orders. If a work center is working on a yellow order and a red order arrives at the work center, it is sufficient that the red order moves to the head of the queue and be processed immediately after the order currently being processed.

Another feature of assigning color codes to work orders is that they provide information about the adequacy of the production lead times that have been established. If the number of red orders is too low, then it is a clear indication that the production lead time is larger than it needs to be. The allowed time is large enough that few, if any, orders are experiencing disruptions that have any consequence. This suggests an opportunity to reduce the planning production lead time used in the DBR system and eventually to reduce the lead time quoted to customers. Conversely, if the number of red orders is large, then it suggests that a large number of orders are experiencing significant disruptions relative to the time allowed for their processing. In this case, the production lead times are too aggressive and need to be increased. It has been our experience that the total number of orders in the red should be around 10 percent. If the percentage of red orders exceeds 15 percent, then we

should consider increasing the size of the buffer. If the percent of red orders is less than 5 percent, then we should consider reducing the size of the time buffers.

Complex Production Environments and a Classification Scheme

Real-life production environments are much more complex than the simple flows used in explaining the DBR system. Even a medium-sized factory has hundreds, and often thousands, of parts and products and has tens, and often hundreds, of different resources. In other words, the detail complexity of production operations is immense. When the focus is on the detail complexity, one is overwhelmed and tends to believe that each operation is unique and there is little that can be transferred in learning from one operation to the next. Typically, production is lumped with the business as a whole and is described in terms of the industry segment to which they belong, such as an auto plant or food and beverage plant. In this section, we present a scheme to organize the production operations based on their product flow characteristics. This classification scheme will bring together elements of detail complexity and dynamic complexity that have an impact on managing the production operations effectively. By managing effectively, we mean deliver the products as promised to customers, while keeping investments in resources and inventory to a minimum. The development of the classification begins with a change of perspective, from a view centered on resources to a view that is centered on product flow.

The Fundamental Elements of the Classification Scheme

Since we are used to the resource centric view of production, we have to learn how to view the same operation in a way that focuses on flow. Such a view is provided by the view of production represented in [Fig. 8-2](#) and [Fig. 8-3](#)—the view of operations from the point of view of the materials. It is a time-oriented description of the manufacturing process. As indicated earlier, the resulting diagram of the production operation is referred to as a PFD.¹² We now explore PFDs in more detail.

Consider the simple case where we have three different raw materials (RM-A, RM-B, RM-C) that are fabricated into three component parts (A, B, C) and that these are assembled together into a finished Product D. This simple production operation has only one finished product.

To construct the PFD, we begin with raw material A (RM-A) at the bottom left-hand side of the diagram ([Fig. 8-11](#)). Each step in the fabrication of the component part A is represented by a box vertically above the box for RM-A. If the fabrication process consists of four steps (this information is typically contained in the routing file or process sheet for component A in the company's ERP system), then we have a series of four boxes in a vertical line as shown in [Fig. 8-11](#). For clarity, inside the box we have designated the process step and the resource used in that step—again a piece of information found in the routing file. The first step is designated A-010 and is performed by resource R1. The second step is A-020 performed by R2, and so on. Similarly, the fabrication process for component B made from RM-B consists of three steps and is represented by the series of three boxes B-010, B-020, and B-030. Finally, component C made from RM-C requires four process steps and is designated by the boxes C-010, C-020, C-030, and C-040.

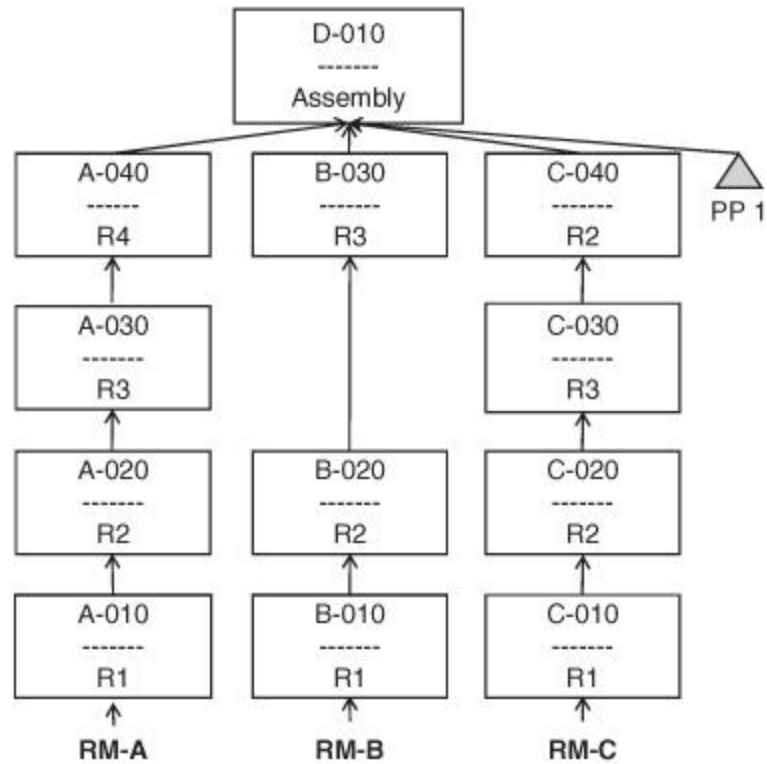


FIGURE 8-11 A detailed product flow diagram for an assembled product.

In just the part of the PFD that we have constructed thus far, two characteristics of production operations (characteristics that make it difficult to manage these operations) stand out. One factor inherent in the PFD is the dependency of operation B-020, for example, on operation B-010. This type of dependency is referred to as material dependency. Simply stated, B-020 cannot be performed unless B-010 has been completed. Every stage in a PFD depends on the preceding stage. If a box in a PFD has an incoming arrow, this indicates *material dependency*. *The material from the box at the base of the arrow is an absolute requirement for the box at the tip of the arrow.* The boxes RM-A, RM-B, and RM-C have no incoming arrows as they are the beginning of this production operation. If we were looking at the entire supply chain, then clearly these boxes would be linked to the suppliers of these materials.

A second form of dependency that is highlighted in a PFD is between steps A-010, B-010, and C-010. *All of these processes require the same resource, R1.* This is an example of the type of dependency referred to as *resource dependency*. If R1 is engaged in step A-010 and there is only one resource R1, then B-010 and C-010 cannot be performed. Another resource dependency can be seen between stages C-020 and C-040. Both require the same resource R2. In addition, R2 will have to complete C-020 and R3 complete C-030 before C-040 can be started.

In [Fig. 8-11](#), we complete the PFD for this simple operation by adding the assembly operation. An assembly operation, by its very nature, requires more than one input material. Just as the arrow from RM-A to A-010 represents the fact that RM-A is an input to the processing step A-010, the arrows from A-040, B-030, and C-040 all converging on box D-010 indicate that all of the components A, B, and C are required to perform this assembly step. If even one of them is missing, the assembly operation cannot proceed. In [Fig. 8-11](#), the arrow from PP1 to D-010 represents the fact that a purchased Product PP1 is required (in addition to parts A, B, and C) to perform operation D-010. We refer to assembly stages as *convergence points* in the PFD—*multiple products/materials are assembled together to make a single product*. A convergence point (a control point) represents a high degree of dependency since all materials represented at the base of the multiple arrows are necessary for this operation to be performed.

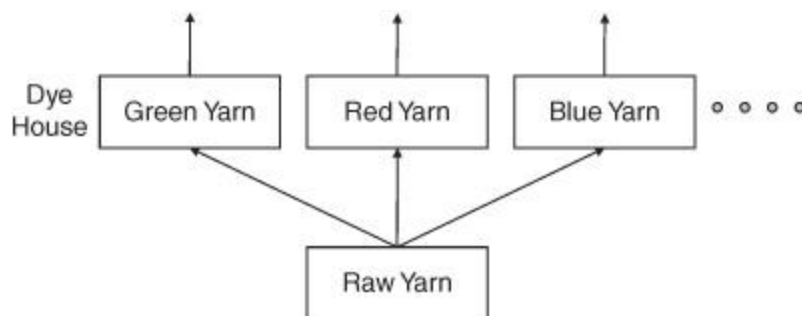


FIGURE 8-12 Product flow diagram illustrating a divergence point.

In addition to the linear and converging flows, there are cases where the flow shows a divergence. Just as convergence is characterized by the coming together of multiple materials into a single product or component, *divergence* (a control point) is characterized by *a single material being transformed into several different output materials*. Consider, for example, a case in the textile industry. [Figure 8-12](#) shows the case of a specific type of yarn being processed at the next stage—the dye house. At the dye house, color is applied to this yarn. We know that for the same yarn, different colors can be applied (red, blue, green, etc.) and we also know that red yarn is a distinct and different product than blue yarn. In the language of the PFD, the dye house is a divergence point—the same input material (untreated yarn) can leave the dye house as any one of a multitude of colored yarns. The divergence point at the dye house shows up in the PFD as a single yarn diverging at the dye house into a multitude of different boxes.

Material dependency, resource dependency, convergence points, and divergence points are the fundamental elements of a PFD. As discussed in the next section, production operations can be classified into families based on which element is the dominant element in the PFD of that particular operation. If divergence is the dominant element, then we have a V-plant. If convergence is the dominant element, then we have an A-plant. If both divergence and convergence exist (and exist at the same stage), then we have a T-plant. If we have neither divergence nor convergence, then we have a simple case of resource contention and the plants are classified as I-plants.

V, A, T, and I Flows—Descriptions and Examples

V-Plants

V-plants are dominated by the presence of divergence points throughout the product flow. The PFD for a plant that exhibits divergence at every step is shown in [Fig. 8-13](#). Notice that this diagram resembles the letter V; hence, the name V-plant. In addition, in most real life V-plants, the different products share common resources at most stages in the process. A steel rolling mill provides a good example of a V-plant. The first step in the process is annealing where the sheets of steel are softened in preparation for rolling. At the rolling operation, a given piece of steel can be rolled into any of a large number of different thicknesses. Rolling represents a divergence point. At each divergence point, the number of distinct products increases. For example, after rolling each of the different thicknesses, steels can be heat-treated to many different products with different strength and hardness characteristics (based on the manner of heat-treating). Each of these steels, now with unique thickness and mechanical properties, can be cut into desired widths at the slitting operation. From just a few varieties of steel coils at the start of the operation, one can end up with thousands of finished products—characterized by thickness, mechanical properties, width, and length.

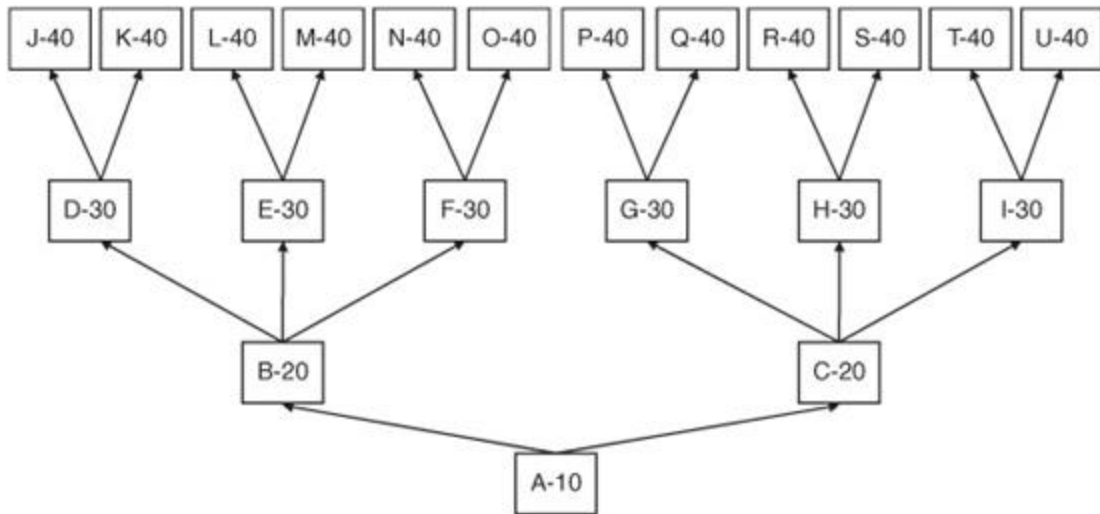


FIGURE 8-13 Product flow diagram illustrating a typical V-plant.

The existence of divergence points gives rise to three primary characteristics of a V-plant regardless of the specific industry or materials.

1. The number of end items is large compared to the number of raw materials. Because divergence points exist throughout the different stages of production, by the time several stages are completed, the number of different products can be very large as can be seen in the rolling mill example.
2. All end items are produced in essentially the same way. All products are processed through the same basic operations—rolling, heat-treating, slitting, etc.
3. The equipment is generally capital-intensive and highly specialized. The evolution into capital-intensive equipment is not difficult to understand. Since every product goes through the same sequence of operations, there are a relatively small number of basic operations performed repeatedly. Because the

focus of improvement under the traditional cost-based system is to reduce the product's direct labor content, the equipment naturally became specialized, high-volume, capital equipment.

The one characteristic that all V-plants share is that despite having high levels of finished goods inventory, there is constant scrambling to meet customer requirements. The capital-intensive nature of the equipment, which typically comes with lengthy setup times and the presence of divergence points, is at the heart of this problem. The lengthy setup times encourage supervisors to increase batch sizes, to minimize setups by combining batches whenever possible, and to produce families of products together. All of these actions, which are consistent with cost-world thinking,¹³ result in a mismatch between customer required priorities and production priorities. In addition, the large production batches cause the production lead times to increase. The result of all of these actions is that lead times are long and unpredictable and this ultimately leads to missed due dates.

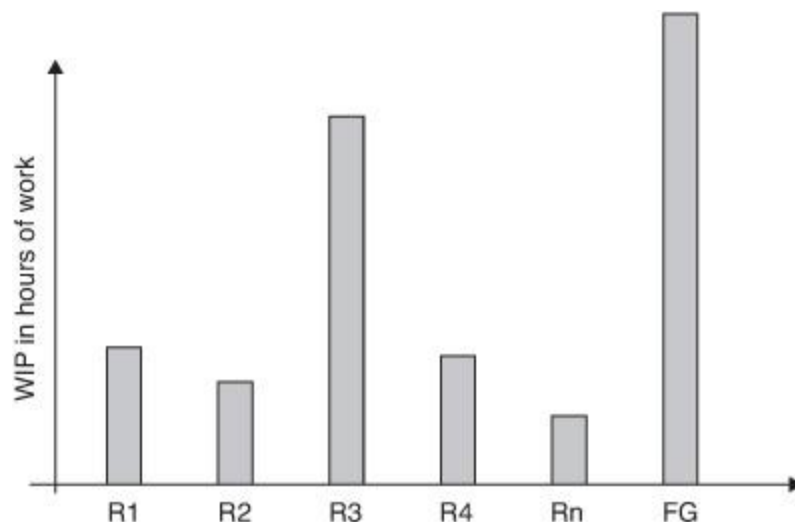


FIGURE 8-14 The WIP profile of each resource (in hours of work for that resource) for a V-plant.

V-plants typically face the following concerns:

1. Finished goods inventory is large.
2. Customer service is poor.
3. Manufacturing managers complain about constantly changing demand.
4. Sales and marketing managers complain about the lack of responsiveness from manufacturing.
5. Interdepartmental conflicts are common within the manufacturing area.

DBR in V-Plants

It is important to recognize first that in almost all cases, there is a considerable effort underway to address the problems faced by a typical V-type plant. Each of the issues is assigned a cause and a solution is either being designed or in implementation. However, the problems persist in most cases. A properly implemented DBR solution will address many of the root cause issues that underlie the V-plant problems and thereby help mitigate most of these problems at the same time. If a capacity constraint exists and these are the only conditions in which the full DBR system would be considered, then identifying which resource is the capacity constraint is the first task. In V-type plants, this is a simple task. Since the resources are involved in the flow of most products, material naturally accumulates in front of the resource with the highest load. The CCR is thus the resource with the largest in-process queue (measured in hours of work for that type of resource). In the case shown in [Fig. 8-14](#), the constraint is resource R3. It is also true that

the personnel in the plant have a common and usually correct knowledge of the constraint. As an aside, it should be noted that the presence of high levels of Finished Goods (usually the largest bank of inventory in the flow) suggests that setups are considered large at many key resources in the operation.

The next key step is to establish the drum. The challenge in most V-plants is the fact that the load placed on a specific resource is significantly influenced by the number of setups that result from this mix. In other words, changing the product mix can change the resource that is most loaded. For example, a textile mill running very large batches of a given color can significantly reduce the total load at the dyeing resources but can cause major problems at the cutting and sewing operations. This is because a single color material will have to be fabricated into apparel of many different sizes and styles and this causes overloads at these work centers. The key to establishing the drum is to find the proper balance between the market demand and the schedule at the constraint that satisfies the requirements for a drum:

1. It satisfies market demand.
2. It maximizes Throughput for the system.
3. It does not create new constraints.

The other factor in designing and implementing a DBR system in a V-type plant that needs special attention is the existence of a large number of divergence points. Each divergence point is a schedule control point and needs to be managed as such. Detailed lists that show the different products that need to be produced and the exact quantity that needs to be produced for each product are required at each divergent point resource.

The schedule control points are material release, constraint(s), divergence points and shipping.

A-Plants

A-plants are characterized by the existence of convergence points wherein a large number of component materials are assembled together into a few end items. The component parts are usually made up of parts that are fabricated in the plant (or other plants/departments in the division) and parts that are purchased from outside vendors. The typical PFD for an A-plant is shown in [Fig. 8-15](#). One characteristic of A-plants, which is different from characteristics of T-plants, is the fact that the *component parts tend to be unique to a single end item*. Several levels of subassemblies may be involved prior to final assembly. Since the overall product flow is convergent rather than divergent, the product flow diagram resembles an inverted V, thus the designation A-plant.

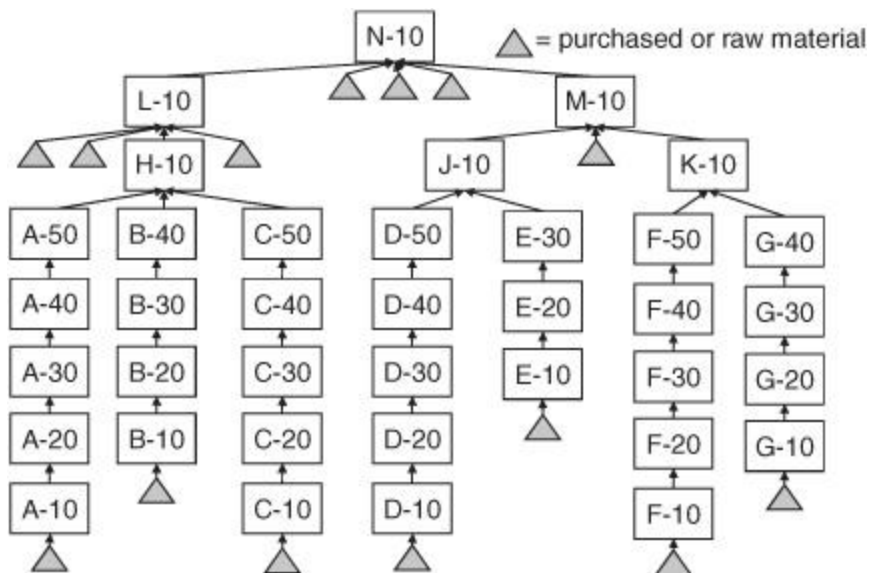


FIGURE 8-15 Product flow diagram for a typical A-plant.

One example of an A-plant is provided by aircraft manufacturing. The PFD contains several thousand components that converge to a single product. The components that arrive at the final assembly plant are themselves major assemblies (jet engines, for example). In addition, the number of distinct aircraft types is quite limited—for example, Boeing has fewer than 10 active models.

The general characteristics shared by A-plants include:

1. Assembly of a large number of manufactured and purchased parts into a relatively small number of end items. Each assembly point represents a decrease in the number of distinct parts and after just a few assembly steps, the number of distinct items drops dramatically.
2. The component parts are unique to specific end items. This is a key feature that distinguishes A-plants from T-plants. Consider aircraft, for example. While every aircraft has engines, the engine for each type of aircraft is unique. The engine for a Boeing 747 is completely different from the engine for a Boeing 777.
3. The production routings for the component parts are highly dissimilar. In the example of the aircraft, the routing for the manufacture of a jet engine blade is nothing like the routing for the manufacture of the compression chamber.
4. The resources and tools used in the manufacturing process tend to be general purpose. In an A-plant, the same resources are used to produce many different parts. Resources are quite flexible, in contrast to the highly specialized equipment in V-plants.

Since the major focus in traditional manufacturing environments is resource utilization and not product flow, it is not surprising that the flow through fabrication and into the finished components is erratic. In fact, the flow through all areas of an A-plant is wave-like, resulting in what is characterized as “feast or famine.” This wave-like flow means that it is highly unlikely that all of the component parts are available when needed at assembly. The missing parts must be tracked down and expedited to assembly. The feast or famine syndrome also creates the perception that bottlenecks “wander.”

The major concerns in an A-plant include:

1. Assembly is constantly complaining of shortages and expediting is a way of life in manufacturing and purchasing.
2. Unplanned overtime is excessive. Resources that were idle during the week suddenly find a wave of material that is needed urgently at assembly in their queue and this results in overtime.
3. Resource utilization is unsatisfactory.
4. Production bottlenecks appear to wander about the plant.
5. The entire operation appears to be out of control.

DBR in A-Plants

Unlike V-plants where the identification of the constraint is straightforward, the identification of the real capacity constraint is not straightforward. This is a direct result of the possibility of product flows for different component parts being different. This can create a situation in which multiple

constraints appear to be present. In addition, the use of large production batches (chosen to reduce unit costs and improve resource efficiency) results in wave-like flow and the constraint appears to wander from one resource to the next. At first sight, it might appear that the resource load information from the computer planning systems would provide a simple means of identifying the constraint, especially since most of these plants have a computerized planning and control system. However, the data are unreliable to the extent that in the author's experience resource load data from the computer system is highly suspect. As explained in Srikanth and Umble (1997), the best way to identify the constraint is to analyze resources that use the most overtime on a regular basis with the parts shortage information (the daily shortage list from assembly). The resource that uses overtime and processes parts regularly on the shortage list must be the constraint.

Two key factors should be considered in setting up the drum in an A-plant. The first is that the assembly (convergent point) operation provides an excellent place to establish the drum. Subordinating everything else to a well-constructed assembly schedule is the easiest way to achieve a good smooth flow through the entire operation. The assembly schedule should be established in such a way as to:

1. Meet market commitments.
2. Be within the constraint's capabilities.
3. Achieve a smooth flow through all of the operation.

The second factor is that the batch sizes that are being used are often too large and should be significantly reduced. Small batches are key to achieving a smooth flow and they should be aggressively reduced. Keep in mind that a batch

is too small only when it creates a capacity constraint due to the increased number of setups that might be caused.

The schedule control points are material release, assembly, shipping, and the physical resource constraint (if one exists).

T-Plants

The critical feature of a T-plant is that the final products are assembled using a number of component parts and these component parts are common to many different end items (in contrast to an A-plant). Because of this sharing of components, the assembly part of the product flow has the structure shown in [Fig. 8-16](#). Note that the number of end items is larger (much larger) than the number of component parts. This creates the sudden explosion of the PFD to create the T-shape. To illustrate the magnitude of this explosion, consider a case where there are six component parts and each part has four variations, giving a total of 24 different components. The number of possible end products is $4 \times 4 \times 4 \times 4 \times 4 \times 4 = 4096!$

Most manufacturers of consumer products are T-plants. Consider the production of personal computers. The basic elements—hard drive, processor, memory, display, etc., are available in a few variations each. For example, the hard drive is available in 40, 60, and 80 G sizes. The processor might be available with speeds of 1.8, 2.0, or 2.4 GHZ. As illustrated above with just a few such variations, the number of distinct computers that the manufacturer produces can be very large indeed.

The characteristics of a T-plant are:

1. A number of common manufactured and purchased parts are assembled together to produce the final product.

2. The component parts are common to many different end items.
3. The production routings for the fabricated component parts are usually quite dissimilar.

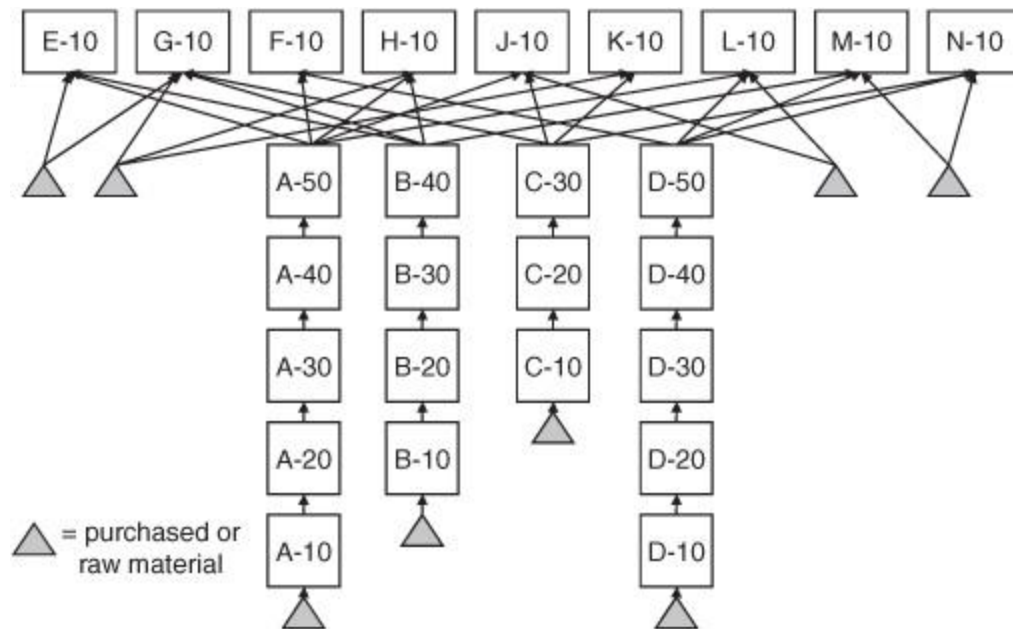


FIGURE 8-16 Product flow diagram for a typical T-plant.

The dominant characteristic of a T-plant is that the assembly point is actually a divergence point. The same component part (80 G hard drive) can be assembled into a very large number of different end units. Unlike a V-plant where the divergence points are spread through the operation, the divergence in a T-plant is concentrated in the assembly area. The impact of this is devastating. We have seen the impact of a simple divergence point in the case of V-plants. In a T-plant, the divergence is assembly and this means that not one but all components are diverted to the wrong product if assembly produces the wrong item. This significantly magnifies the impact and spreads through the whole system like wildfire. This is illustrated by the simple case shown in [Fig. 8-17](#) involving four component parts A, B,

C, and D, and four assembled Products E, F, G, and H. The arrows show how the products are made and the figure indicates the inventory available for each part. Now suppose that an order for 100 parts of Product E is due to be assembled and shipped. The assembly of Product E requires 100 units of part A and 100 units of part B and is next on the assembly schedule. However, as shown in [Fig. 8-17](#), part A has zero inventory. An expediter will have to be dispatched to expedite 100 units of part A. In the meantime, the assembly operation is going to be idle. However, it is possible to make 100 units of Product H, which requires part B and part C. In most cases, the assembly will not be left idle. Product H will be produced, since it is an active part and might very well have an order due next week. Note that this action consumes the available stock of part B, while at the same time creating finished inventory of Product H. Alternately, Product E is behind schedule while Product H is ahead of schedule. However, the real damage is revealed when part A finally arrives at the assembly area. It is still not possible to assemble Product E because we are now short of part B that was consumed in the production of Product H.

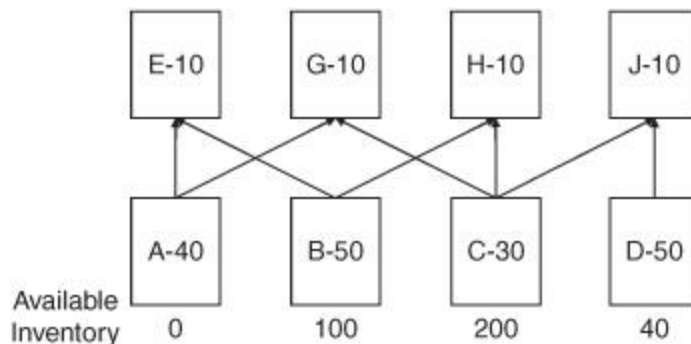


FIGURE 8-17 Example of the phenomena of “stealing.”

The concerns or issues that are shared by T-plants in general are:

1. Large finished goods and component inventories.

2. Poor due date performance (30 to 40 percent of the orders early and 30 to 40 percent of the orders late).
3. Excessive fabrication lead times.
4. Unsatisfactory resource utilization in fabrication.
5. Fabrication and assembly act as separate unsynchronized plants.

DBR in T-Plants

In a T-plant, we have two situations. The most common situation is that most T-plants tend to belong to the MTS environment. Typically, buffer stocks are maintained at both the component level (just prior to assembly) and at the finished goods level. In this case, there are no real constraints (see discussion of MTS in [Chapter 10](#)) and the proper system to implement is the S-DBR system discussed in [Chapter 9](#).

If this is not the case and there are capacity constraints, then the key factor is that the assembly operation must be managed properly. As long as stealing occurs at assembly, T-plants will be chaotic and flow will be difficult to manage. However, once stealing is eliminated through tight control of the assembly operations, then a T-plant becomes an A-plant and the DBR system discussion in A-plants should be followed.

The schedule control points are materials release, divergence, convergence, and the physical resource constraint (if one exists).

I-Plants

I-plants are the simplest of the production flows. The major issue with I-plants is the sharing of resources between the different products. Each product follows the same sequence of operations. There is little or no assembly and there are no divergence points.

The characteristics of an I-plant are:

1. All parts have similar routings.
2. Resources are shared between different parts, while raw materials are not.
3. There is very little assembly involved.

The typical I-plant product flow is shown in [Fig. 8-18](#) and the shape makes the name obvious.

I-plants are the simplest of plants to manage. Nevertheless, traditional focus on resource utilization results in the use of production batches that are much larger than required to maintain a smooth flow. As a result, WIP piles can be created and the wave-like flow of A-plants can be observed. Consequently, I-plants have the following concerns/issues:

1. Low due date performance.
2. High WIP inventories.
3. Level of output below theoretical line rates.

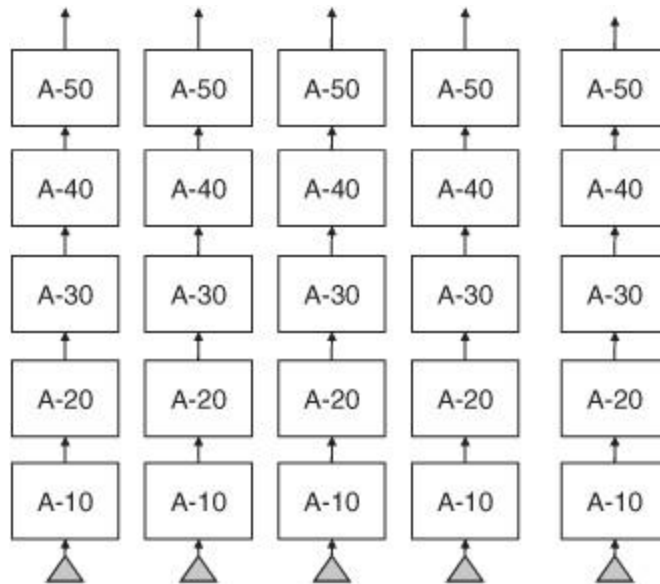


FIGURE 8-18 Product flow diagram for a typical I-plant.

DBR in I-Plants

I-plants are straightforward to manage from a product flow standpoint. The DBR system as described in the previous sections can be designed and implemented with little complication. Identification of the constraint is simple—all personnel will be aware of this resource and the inventory buildup should confirm the section of this resource. Simple steps to improve productive use of this resource (see the section on Step 2—exploiting the constraint) should be followed by the implementation of the DBR system.

Most academic research has been conducted on I-plants (primarily on lines of 10 or less work centers) as indicated in [Chapter 7](#). It is by far the simplest to simulate and study. In contrast, most plants are V, A, T, or combinations of these structures.

Summary

This chapter covered the basic terminology and concepts related to the TOC production solution. As such, it provides the foundation for a deeper understanding of DBR in an MTO environment, S-DBR in an MTA environment, and supply chains linking manufacturing to the downstream links. The various types of buffers are defined and illustrated, as are the various types of plants with their control points. A discussion of implementing DBR in each environment is provided.

References

- Blackstone, J. H. 2008. *APICS Dictionary*. 12th ed. Alexandria, VA: APICS.
- Ford, H. 1928. *Today and Tomorrow*. Garden City, NY: Garden City Publishing,
- Goldratt, E. M. 1990a. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Crotonon-Hudson, NY: North River Press.
- Goldratt, E. M. 1990b. *What's This Thing Called Theory of Constraints and How Should It Be Implemented?* Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 2003. *Production: The TOC Way*. Rev. ed. Great Barrington, MA: North River Press.
- Goldratt, E. M. 2009. "Standing on the shoulders of giants" *The Manufacturer* June. http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants. (accessed February 4, 2010).

Goldratt, E. M. and Cox, J. 1984. *The Goal: Excellence in Manufacturing*. Croton-on-Hudson, NY: North River Press.

Schragenheim, E. and Dettmer, H. W. 2001. *Manufacturing at Warp Speed*. Boca Raton, FL: St. Lucie Press.

Senge, P. M. 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. New York: Doubleday Currency.

Srikanth, M. and Umble, M. 1997. *Synchronous Management: Profit-Based Manufacturing for the 21st Century*. Vols. 1 and 2. Guilford, CT: Spectrum Publishing Company.

Sugimori, Y., Kusunoki, K., Cho, F., and, Uchikawa, S. 1977. "Toyota production system and Kanban system materialization of just-in-time and respect-for-human system." *International Journal of Production Research*, 15(6):553-564.

Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/default.asp?page=dictionary>.

About the Author

Dr. Mokshagundam (Shri) L. Srikanth obtained his PhD in physics from Boston University. After a brief tenure as Associate Professor at Boston University, he joined Dr. Eli Goldratt in 1979. He is a partner in the Goldratt Group, an international organization headed by Dr. Eli Goldratt and dedicated to helping organizations and individuals achieve breakthrough improvements through the creation and

dissemination of new knowledge. He is currently head of Goldratt Schools for North America.

He has nearly three decades of experience with industrial enterprises and ways to improve their performance. Dr. Srikanth was a Senior Director of the Center for e-Business Excellence at i2 Technologies. Prior to this position, he was a Director in i2's Product Management group. Before joining i2 Technologies, he was cofounder and managing principal of Spectrum Management Group.

Dr. Srikanth has helped companies improve delivery performance, reduce lead times, and reduce investment in inventories and resources. His experience covers a broad cross-section of industries including aerospace and defense, automotive, furniture, textiles, consumer, and industrial products. Companies range from Fortune 100 companies such as General Electric, Ford, General Motors, and United Technologies to small family-owned organizations.

Dr. Srikanth has authored several books including *Regaining Competitiveness: Putting 'The Goal' to Work*, with Harold E. A. Cavallaro, 2nd Revised Edition (North River Press, 1993); *Synchronous Manufacturing: Principles for World Class Excellence*, with Professor Michael Umble (Southwestern Publishing, 1991); *Measurements for Effective Decision Making*, with Scott A. Robertson, (Spectrum Publishing Company, 1995); and *Synchronous Management: Principles for Profit-Based Manufacturing for the 21st Century*, Vols. 1 and 2, with Professor Michael Umble (Spectrum, 1997). He is a contributor to Srinivasan, Mandyam, *Streamlined—Principles for Building and Managing a Lean Supply Chain* (Cengage Learning, 2004).

CHAPTER 9

From DBR to Simplified-DBR for Make-to-Order

Eli Schragenheim

Introduction

Drum-Buffer-Rope (DBR) is the name given by Dr. Eli Goldratt to a simple and effective production planning method. The root of the name is based on the analogy of the scouts tour described in *The Goal* (Goldratt and Cox, 1984, [Chapters 13-15](#)). DBR was at the time the cornerstone of the *Theory of Constraints* (TOC) and continued to be the best known application of the theory until the appearance of *Critical Chain* (Goldratt, 1997), outlining the concepts for planning projects.

Simplified Drum-Buffer-Rope (S-DBR) is a variation on the original DBR methodology. It was suggested by Schragenheim and Dettmer (2000) in *Manufacturing at Warp Speed* as a valid, simplified replacement especially suited when the implementation has to use the common material requirements planning (MRP)/enterprise resource planning (ERP) software. Since then, the basic principles of S-DBR were adopted by Dr. Goldratt. Important improvements were added and dedicated software for S-DBR has been developed by Inherent Simplicity Ltd. under the close supervision of Dr. Goldratt. S-DBR has now replaced the older DBR as the preferred planning method

with one exception, which will be explained later in this chapter.

Another important realization concerning production planning has to be mentioned. Both S-DBR and DBR were assuming a make-to-order (MTO) environment. During the rethinking of the TOC-focused planning methodology, it was recognized that the make-to-stock (MTS) production environment should be based on different principles. The author dedicates [Chapter 10](#) to MTS, or rather to make-to-availability (MTA)¹ environments, to emphasize the clear distinction.

Another comment should be made. While DBR and S-DBR are planning methods, they are not stand-alone methods. *Buffer Management* (BM), the TOC control mechanism, should be viewed as inseparable from the planning method. Thus, [Chapters 9](#) and [10](#) deal with both the DBR/S-DBR planning as well as with BM as an absolutely necessary part of both planning methodologies.

The purpose of this chapter is to explain the S-DBR/BM concepts, logic, and procedures through the development of the ideas over time. Thus, the emphasis is on the historical development, which is critical to the full realization of the continual paradigm shift we have gone through during the last 25 years since the introduction of DBR.

A Historical Background and Perspective

In the mid-1980s, DBR represented a huge advancement in providing a robust plan for the production floor. DBR was developed as a major departure from the concept, created by its own developer, of very sophisticated and detailed planning of the shop floor. In the late 1970s through the first half of the 1980s, Dr. Eliyahu M. Goldratt led a software company, Creative Output Ltd., in developing a sophisticated program called OPT[®] (*Optimized Technology*)

to plan manufacturing orders in great detail for any kind of production shop floor. OPT[®] was a true advanced planning and scheduling (APS) program even though the term was coined years later. At the time, the name given to such programs was “finite-capacity scheduling system” and that name hinted at a contrast with the MRP II programs of the time, which were known as “infinite-capacity scheduling systems.”

DBR came as an antithesis to the OPT[®] concept and it came from the author of OPT[®] development—Dr. Goldratt himself. Instead of ultra-sophistication in trying to solve a complicated net of links between the processing steps and resources, of which several might have limited capacity (bottlenecks), a vastly simplified concept emerged: in any chain, there is *one link* that is the weakest. That link determines the strength of the whole chain; thus, detailed planning of that specific link should be the kernel of the overall production plan. The name given to the core planning—scheduling the one bottleneck ensuring its smooth and effective utilization—was the *drum*. The resulting understanding was that *the bottleneck is the only resource whose efficiency really counts*. However, planning the bottleneck does not ensure that the plan would be executed as is. *Murphy, the symbol for everything that might go wrong*, could mess things up and the bottleneck might face a situation where it has to stop processing because parts are missing. Instead of sophisticated synchronization of all resources, the concept of providing a *buffer to protect the bottleneck from being starved* had emerged. This buffer is not made of stock—it is a *time buffer*. The idea was to *release the materials for the bottleneck exactly a time-buffer length before the bottleneck is supposed to begin work on the job, giving all the required resources enough time to let the parts reach the bottleneck before the scheduled time*. This concept of

the buffer as time—supporting the timely arrival of the parts rather than parts sitting in front of the bottleneck—was a key in understanding the paradigm shift that lies in the change from great sophistication to simplicity. It consists of understanding that buffers are necessary to deal with uncertainty, and that in order to protect a schedule, which is built of time-based instructions, we need to use time as protection. The time buffer meant that even when Murphy messes things up, the expectation is that the parts will reach the bottleneck on time in the vast majority of cases. Of course, specifying long enough time to cover for Murphy meant that in most cases the parts would arrive too early to the bottleneck and simply sit there. So, it looks like a buffer of inventory, but actually the real protector against starvation of the constraint is the time provided for parts to go through the route to the bottleneck.

The term “bottleneck” was the key term in the OPT[®] days, and even when the DBR methodology was developed together with the famous book *The Goal* (Goldratt and Cox, 1984), the terminology was based on bottlenecks. It is always important and enlightening to have a historical perspective of the development of such major managerial approaches as TOC. At that time, in 1984, the much more generic term *constraint* was not yet coined.

OPT[®] registered trademark of Scheduling Technologies Group Limited, Hounslow, U.K.

The important insight, partially acknowledged in the OPT[®] days, but becoming clearer later, is:

As complex as the production shop floor may be, the performance of the shop as a whole is impacted by a single work center, which determines both the response time and the maximum potential output of the floor.

Is there really only *one* capacity constraint (called CCR—*capacity constrained resource*), or could there be two? Well, technically, it is possible to have two, but assuming we speak about interactive resources (one feeds the other) being driven to their limits, then the performance of the shop is doomed to be unstable and even erratic because of the statistical fluctuation that inevitably occurs between dependent resources.

This chapter is not focused on DBR, but on S-DBR and the transition of the understanding that paved the way from DBR to S-DBR. We just stated one transition from OPT[®] to DBR and the main way is still ahead of us. Before we proceed, let's fully understand three different aspects of the TOC approach. Each is material in understanding the development from DBR to S-DBR and the internal logic of S-DBR.

Three Views on Operations Planning and Execution

The basic TOC philosophy was first expressed by the *Five Focusing Steps* (5FS), which already explain the logic of the TOC production planning and its related BM control. The second viewpoint recognizes the difference of defining the rules behind planning in a world with a significant amount of uncertainty (planning with uncertainty) versus planning to optimize in a deterministic world. At the time of the execution whatever is dictated in planning lays out the objectives and the resulting actions. But then, there is a need to define the rules for the decision-making required to deal with the impact of “Murphy” in executing the plan. It is fascinating to realize that defining the rules for planning and execution lead to the S-DBR planning rules and the role of BM in leading the execution decision-making.

The third viewpoint looks at the achievements of Henry Ford and Taiichi Ohno and their focus on flow as the central

objective of operations. It seems that even that viewpoint fully supports the TOC methodology for production planning and execution. The three different aspects bring together a better understanding of the methods and how to match them to different environments.

The Five-Focusing Steps (5FS)

The concept of the 5FS³ was developed in 1985 just as internal knowledge transfer within Goldratt's company, Creative Output signaled the emergence of the comprehensive managerial approach of TOC. It is the first time the term *constraint* had replaced the older concept of a *bottleneck*.

The importance of the 5FS (Goldratt, 1990b, 7) is that they define the rules for a “well-behaving organization.” The first three steps define the state of the short-term:

1. *Identify* the system constraints.
2. Decide how to *exploit* the system's constraint.
3. *Subordinate* everything else to the above decision.

The longer-term steps give an umbrella for developing the scheme for growth coupled with stability:

1. *Elevate* the system's constraint.
2. Go back to Step 1. Warning: Beware of inertia.

For a better understanding of the DBR methodology and the transition to S-DBR, only the first three steps have a direct impact. Beginning in 1985, the three steps were extensively used to explain DBR thinking—even though the DBR rules preceded the 5FS. The first three steps are prominent in explaining the shift to S-DBR.

The Critical Distinction between Planning and Execution

The Appropriate Rules of Planning

The role of planning is to synchronize the system in a way that would enable achieving its objectives. Many times, the planning affects objectives by identifying what is realistic to achieve and what is not. Planning is viewed as the higher-level decision making, while the execution is viewed as just having to follow the planning.

There are two main difficulties for any kind of planning. One is the internal complexity in synchronizing many different variables. The other is dealing with uncertainty. The main problem in dealing with uncertainty is that planning decisions are made ahead of time and most decisions are converted to specific actions. This time difference between planning and execution allows Murphy to mess things up to a degree that the planning cannot be executed as is. The situation where in the execution phase it seems impossible, or not worthwhile, to follow the planning not only causes problems in achieving the system objectives but also generates tension between the planners and the people in charge of the execution.

Viewing the DBR methodology versus OPT[®] might shed light on the way TOC treats the planning rules. Later, we will look at the resulting insights regarding the impact of TOC on decision rules in execution.

OPT[®]⁴ was all about planning. It planned all the perceived bottlenecks in detail under finite-capacity and then went on planning the rest of the shop floor where all the non-bottlenecks were scheduled under the infinite-capacity assumption in a similar way to MRP. The hidden assumption was that there was no need to make any significant decisions at the execution phase—just follow the schedule.

If Murphy messed things up, then running OPT[®] again was the reasonable option.

DBR is a planning algorithm that is much less detailed than OPT[®]. Only one constraint is scheduled in detail.⁵ All the rest of the resources are not given any schedule.⁶ However, the material release was scheduled in detail with the notion that the schedule for the material release meant: *Do not release before!*

Our current understanding is that having good planning means that, in most cases, the plan is eventually executed without changes and it draws good performance from the shop floor as a whole. Any instruction that is included in the plan but is not absolutely necessary to be taken at the time of the planning endangers the sustainability of the whole plan. The rules for what should be included in good planning are:

1. Any instruction where any deviation might disrupt achieving the objectives.
2. Such instructions must be protected from Murphy. Buffers have to be included in the plan protecting the ability to carry out the instructions.
3. Nothing else should be included in the planning.

The DBR methodology has clearly defined the critical points in the product structure that must be planned carefully. Three major control points are:

1. The *due dates for all orders* after careful validation that these dates are quite safe.
2. The detailed schedule of the CCR.
3. The schedule for the *material release*.

The criticality of the first control point is self-evident—we should not commit to dates we cannot meet. The second one is simply the essence of Step 2 in the 5FS—exploiting the system constraint. The criticality of the third control point is not self-evident. We usually see in many environments, particularly in manufacturing, that there is a lot of *work-in-progress* (WIP) that just sits and waits for resources. The immediate cause is the release of work too early to the shop because the first resources are available. The assumption is that the earlier work is released and starts, the higher the probability of finishing the work on time. However, once the first resources finish processing those orders they simply join the queue for the subsequent resources. The damage of having too much work without a clear and rigorous priority mechanism is enormous. While the resources upstream might be looking for work, some of the other resources might be flooded with work. When this happens, the resource under pressure tries hard to optimize its own efficiency, often at the expense of orders that are truly urgent. Actually, in most cases, the operators do not have any idea what is urgent and what is not. Many manufacturing orders are comprised of large batches. When the manufacturing orders are large, they often contain urgent customer orders and much less urgent stock orders all in one manufacturing order. Thus, too many manufacturing orders have a certain quantity that is very urgent to customers and some other quantity that is not. The loaded resource cannot do all of the manufacturing orders at the same time. Therefore, while a work center is working on a large manufacturing order with several urgent customer orders buried within it, other manufacturing orders that may also have urgent customer orders wait their turn.

The *rope* in the DBR planning methodology is *the mechanism to ensure release of only orders that are soon required by the detailed schedules of the CCR and shipping*

buffers. This mechanism also forces the minimization of batching. The rope ensures against work that is not truly required being released to the shop.

The Implications for the Execution Phase

We have shown the general idea behind “minimum planning.” Let’s now describe the execution decision-making rules. When planning is not detailed, much more is left to the execution phase.

Including buffers in the planning has a special meaning for the people in the execution phase. The local objective in execution is to be able to execute the critical planning instructions. The state of the buffer is an excellent signal of whether things are going according to plan.

BM is the control mechanism on the progress of executing the plan. First, let’s introduce my definition of the term *control*:

Control is a proactive mechanism to handle uncertainty by monitoring information that points to a threatening situation and taking corrective actions accordingly.

The definition makes it clear that any control system is targeted to identify the actual emergence of a known threat, and it clearly belongs to the execution phase. It must have the most current and accurate information that the execution people need to carry out their jobs.

The need in the execution phase is to validate that everything is ready on time for the next critical directive based on our planning. The obvious possible threat is either being late to the CCR and thus starving the constraint or being late to the due date of the whole manufacturing order and thus making the customer order late. These two areas

are protected by time buffers according to the DBR methodology.

Let's define the *state of the buffer* as the percentage of the time the buffer has already used (the time that has passed since the start of the time buffer). When the state of the buffer is less than 33 percent, we call the state (region or zone) *green*. When it is between 33 and 67 percent, it is the *yellow* state, and above 67 percent it is the *red* state.

A status of red means less than one-third of the original buffer remains and thus it is now priority one to flow the order to its destination (either the CCR or shipping).

Thus, the decision rules for the execution phase are based on the status of the buffers. BM imposes one clear set of priorities and does not tolerate any others. Thus the buffer status of any order can be checked and according to the resulting priorities every resource is able to decide what to do next. Following the BM priorities yields the highest probability that we will ship everything on time and utilize the CCR to its planned exploitation level.

The viewpoint of minimum planning requires the extra emphasis on a priority scheme for the execution phase. It fully supports the move from the excessive planning in OPT[®] to the leaner planning of DBR, but with the addition of BM as the execution aid to achieve the objective of reliable due date performance coupled with good exploitation of the CCR. Later in this chapter, it will be shown how this view also supports S-DBR.

Concentrating on the Flow

The third viewpoint on operations comes from Goldratt's (2009) article, "Standing on the Shoulders of Giants." The flow concepts are attributed to both Henry Ford and Dr. Taiichi Ohno and highlight the TOC approach to planning and execution in manufacturing.

Goldratt⁷ (2009, 3) has verbalized the *four concepts* that lie behind the work and achievements of both Ford and Ohno:

1. Improving flow (or, equivalently, lead time) is a primary objective of operations.
 2. This primary objective should be translated into a practical mechanism that guides the operation when not to produce (prevents overproduction).
 3. Local efficiencies must be abolished.
 4. A focusing process to balance flow must be in place.
-

Certainly, the four concepts for flow link the Lean concept with TOC and particularly DBR (actually, it is more attuned to S-DBR as will be explained later in this chapter). We certainly do like to get faster flow throughout the shop floor. Moreover, the rope is just another tool to prevent any overproduction. Of course, the main point is how to distinguish between overproduction and what should be produced.

The fourth concept is interesting as it can be interpreted as applying both to the immediate state as well as to longer periods. For the immediate situation, it fully supports the idea of giving higher priority to orders that seem “almost late,” thus enabling faster flow for the urgent orders. We still need to develop a more global view on how to focus the efforts on improving the flow in the longer-term period of time.

Challenging the Traditional DBR Methodology

When DBR first appeared in 1984 (*The Goal*), it signalled a departure from the very detailed production planning

process like OPT[®] as well as a contrast to MRP. It was much later that we learned that when planning is minimal then execution gets more responsibility and it needs better guidelines for decisions. BM was mentioned first in Goldratt and Fox (1986), *The Race*. It took additional time for Goldratt himself and other researchers to define clearly the linkages of BM to DBR and additional time for practitioners to understand it fully. BM is a necessary condition for DBR to work effectively.

The three viewpoints poses various questions regarding the central role of an internal capacity constraint. Let's first summarize the claims and then inquire into each of them more deeply.

1. From the *5FS perspective*, there are some questions that are at the core of the challenge:
 - a. Is the proper *strategic constraint* an internal resource? Should the capacity of an internal resource be the constraint of the whole organization?
 - b. Suppose we do have a real capacity constraint in the shop, isn't the market demand a constraint as well? If so, do we have interactive constraints and how do we handle them? In other words, how do we exploit both market and capacity constraints?
2. From the *minimum planning perspective*, the critical question is whether the detailed schedule of the capacity constraint is truly necessary? What would be the damage if the sequence of the constraint would not be followed as is? Do we always lose capacity in such a case?
3. From the *flow perspective*, the challenge is the emphasis on the CCR buffer because from the overall flow it looks like a disruption to the flow. The trigger to the flow is certainly a customer order. Do we need to

create an artificial time delay at the CCR? Is it something that improves the flow or is it a blockage of the flow?

What Should the Strategic Constraint Be?

A worthy capacity constraint for being the strategic constraint is a resource whose capacity is very difficult to elevate. The difficulty might be that it is very expensive, but it could also be that enlarging the capacity is a large project because the ramifications are very substantial for all the functions within the company. Think about a basic steel company where a huge furnace is the most obvious capacity constraint. To build another furnace is a multimillion investment and it takes several years. Then, as building another furnace adds much more than a mere 2 to 3 percent to capacity (in many cases it doubles the capacity), many additional workers are required, not just for the furnace, but also because the new furnace induces elevating most of the other equipment as well.

This is probably an extreme case where the difficulty in elevation is very clear. Even in such a case, a clever CEO might find alternatives to bypass the limitation of the existing furnace by buying basic steel from other manufacturers who do not have a good market for the capacity they have. However, even in the case where the market potential is far larger than the limited capacity of the furnace, the market demand could still be an active constraint because gaining more market demand would improve the performance of the company by allowing producing and selling more of the highly profitable products instead of the less lucrative ones.

Two characteristics of the market demand make it the major practical constraint in the vast majority of cases:

1. Clients do not like to be subordinated to an internal constraint of a supplier of products or services. In most cases, the clients have an alternative supplier, and if that one offers better service, then the clients might choose to move to that supplier. Once this is done, then the company no longer has a capacity constraint.
2. When the potential is far larger than the internal capacity, then the organization can find ways to increase its throughput even without elevating the internal constraint. One obvious way is by increasing the price. Another is to concentrate on the more profitable niches of the market demand.

However, if the market demand is the system constraint, how can an internal resource be a capacity constraint?

The claim is that it is perfectly possible to have both the market demand and the limited capacity of a specific resource as interactive constraints.

Having both the market and one CCR is quite a good match. The necessary assumption is that proper exploitation of both constraints could leave just enough protective capacity on the CCR to ensure that whatever commitment is given to the market will be met. In this manner, the market constraint is given the higher consideration without neglecting the capacity limitation of the CCR.

This is actually the way to handle any situation of interactive constraints: Decide which one is the major or primary one and ensure that the other or secondary constraint would be somewhat less loaded (take less commitments on that one).

Note that a constraint can be defined as anything that cannot subordinate to another constraint and thus cannot be ignored. This definition fits the reality of most CCRs. The market is still the major or primary constraint, but the

capacity limitation of the CCR or secondary one cannot be ignored; thus, fewer commitments from the market should be placed on the secondary constraint.

If the market demand is the major constraint and we see the need to have some protective capacity provided on the CCR, then there is no need to have a special CCR buffer that would protect the sequence of the production orders on the CCR. We still need to carefully monitor the load on the CCR, but not necessarily to schedule the CCR in detail. Of course, we'll expand on this point later when we describe the S-DBR process.

How Is the Planning and Execution Viewpoint Addressing the Issue of Scheduling and Buffering the CCR?

This viewpoint requires validating that whatever is included in the planning is a must. Thus, the question is, "Do we need to schedule the CCR?" Is detailed scheduling the only way to ensure a good enough exploitation of the system constraints?

Once we recognize that even the CCR has to subordinate to the commitments made to the market, then we have to conclude that scheduling the CCR in detail is not required in most cases (later we deal with the one exception). This means also that the CCR buffer, used to protect the schedule of the CCR, is not required and the only buffer that is truly necessary is the one aimed to protect the commitment to the market. The CCR should prioritize its sequencing decisions according to BM at the shipping buffer. However, the load on the CCR still requires monitoring. The insight here is that there is a difference between monitoring the load on a CCR resource and dictating a sequence on that resource.

How Does Refraining From a Detailed Schedule of the CCR Affect the Execution?

The traditional DBR requires three different buffers: the constraint, the shipping, and the assembly buffers (detailed in Goldratt, 1990), but if one concentrates just on the due dates of the firm orders at hand, then only one buffer, the shipping buffer (now called the *production buffer* in S-DBR), covers the whole production time from material release until order completion, is required. Following the green-yellow-red buffer priorities⁸ as they emerge from the use of having one-time buffer per order is much simpler than deciding between a red assembly buffer versus a red CCR buffer or a red shipping buffer.

What Does the Emphasis on Flow Add to the Challenge to Traditional DBR?

This view is fully concentrated on the trigger of the flow—the customer order. The point of the flow is to be able to commit to the client as fast as possible. With this as the main objective, then the challenge to DBR is obvious: Do we really need the constraint buffer or is it a disruption of the flow? After all, the constraint buffer initiates early release of the parts, so on average they reach the constraint and then wait for their schedule to be processed. Having that planned waiting time at the CCR is a disruption to the flow.

It is obvious that there is a need to choke the release to only what is truly required now. This would prevent the tendency in traditional DBR to release certain orders very early in order to exploit the CCR's capacity.⁹ Looking on the CCR as “disruptive to the flow” also highlights the need to be able to quickly elevate the CCR capacity whenever necessary, because even the unplanned wait time at the CCR could be significant due to its relative lack of capacity and any wait time represents a certain disruption to the flow. Of course, this is not always practical, but the basic

thinking is right. We now recognize *in TOC that the ultimate goal is to both grow constantly and remain stable at all times*. Thus, the constraint should not be the capacity of a resource that can be easily elevated because whenever such a resource cannot fully subordinate to the demand (forces too long wait time, thus blocking the flow) it should be elevated. The underlying assumption is that the financial worth of the additional demand that would not be possible to maintain with such a blockage to the flow is worth more than the capacity of a common resource that is easy to elevate.

Outlining the Direction of the Solution

Challenging the DBR procedure of finite-capacity scheduling of the CCR does not mean we are looking for something drastically new. Actually, most of the wisdom included in the original solution is still intact. The most critical insight in DBR, which we have already mentioned, is worth mentioning again:

As complex as the production shop floor may be, the performance of the shop as a whole is impacted mainly by a single work center, which determines both the response time and the maximum potential output of the floor.

This insight is relevant for S-DBR as well, even though the CCR is not scheduled in detail and does not have a specific time buffer. The term “weakest link” is perhaps more appropriate than CCR because the weakest link is not always a constraint. Nevertheless, it is always something to monitor because both the potential maximum output and the possible response time are impacted by it. Thus, the weakest link could be used to signal the sales department

when additional efforts would be most beneficial and when more care should be given to the quoted delivery time given to potential customers.

The Main Ingredients of the Solution

The solution described here refers clearly to make-to-order (MTO). Another chapter is dedicated to make-to-stock (MTS), or rather make-to-availability (MTA).

S-DBR is targeted at the very short-term. Capacity planning for medium- and long-term is not included within the S-DBR methodology, even though certain information could be extracted from the S-DBR and BM that could support longer-term plans.

Regular short-term planning concentrates on:

1. When to quote the due date for production completion. The underlying assumption is that the due date has to be reliable (safe).
2. When to release the materials.

Two critical tools are required for the planning:

1. The time buffer (production buffer) to be assigned to a manufacturing order for a particular product.
2. The load control on one resource. It is possible though to extend the load control to several resources, but only one of them is truly material in dictating the due date and the material release dates.[10](#)

There is yet another piece of information that is important—the standard lead time for the product in the relevant market. It is important because the plan will not always dictate the earliest date that the algorithm based on the load control and the time buffer would come up with. For

example, suppose the standard lead time in the industry is four weeks, but because sales are, at the moment, rather low you can deliver an order in just one week. Would you offer one-week delivery? Well, if the customer is willing to pay a high markup you might agree. Otherwise, this seems like a marketing mistake. First, you might give the impression that you are so pressed for money that you are ready to do anything, even at the expense of quality, just to get the order. Moreover, if you get the order, the future expectation might be that you can always complete the order in a week and when you refuse the request, it is because you don't respect your client.

Thus, the standard lead time is a reference and whenever feasible this should be the date to quote for the regular price and with the promise of absolutely guaranteed on-time completion. Of course, in off-peak periods a company could use the advantage of shorter lead time. Generally speaking, Marketing and Sales should set a clear logical approach to quoting lead times.

The Time Buffer

At the time *Manufacturing at Warp Speed* (Schrageheim and Dettmer, 2000) was written, the shipping buffer was understood as the shortest time we could safely commit to deliver. For instance, if a certain order could be safely delivered in 12 days, then the order had a time buffer of 12 days.

Additional insight from Goldratt led us to distinguish between two different periods that comprise the shortest safe time from order receipt¹¹ to order completion.

1. *The time the order has to wait in queue until the signal from the rope to be released to the shop floor. This prerelease queue time depends on the prior work*

that the CCR has to do. Assuming the CCR has quite a lot to do, then releasing the order immediately does not add any benefit and could cause damage by creating confusion of priorities.

2. *A liberal estimation of the time it takes from order release until completion (the production buffer [12](#))*. As we are going to consider the current queue for the CCR and possibly delay the actual release of the materials for that order, we do not expect a peak load within the shop floor itself. When a peak load happens, then new orders would need to wait longer before being released to the floor, thus decoupling the flow in the shop from the natural pace of the CCR.

The time buffer mentioned in Item 2 is now called the *production buffer*, as it describes the flow time through the shop under regular load. The production buffer does not include transportation or in-transit time to the client.

The issue of delivery date to the client requires a short discussion. The transportation time is an issue only when it is a significant part of the lead time. The question is whether the commitment of the producer includes transportation. In other words, is the transportation time part of the production planning or part of the customer's planning?

Suppose the producer takes responsibility until the goods reach the customer. Then, the production planning should have a due date for completing the production and then have the final delivery date where the transportation time (and possible fluctuations in it) is considered.

[Figure 9-1](#) shows time elements in lead time, but note that from now on we'll treat the completion of the production as the due date to which we refer.

How should the production buffer be determined? In implementing S-DBR in traditional planning and control

environments, the usual recommendation is to cut the current production lead time by half. The rationale is that by eliminating the large batching and huge WIP levels, the main disruption (waiting in queues at each work center) to the flow has been vastly reduced.¹³ Take into account that the net processing time is just a fraction of the production lead time and you realize that by cutting the wait time to half, the total production lead time is cut by half. The priority system of BM supports very high reliability within that time. Thus, cutting the standard production lead time by half is a good initial production time buffer.¹⁴

Noting the standard lead time of the industry is a good gauge for a test measurement and the time buffer to be used should not be longer than half of that number. In most cases, this reduction is not only possible but often the production time buffer can be cut even further. These further cuts should be done only after first implementing S-DBR with buffers that are equal to half the current lead times. After the shop floor has stabilized, the further reduction of the production buffers is achieved through BM with the focus of improving the flow. Recall improving flow is the main mission of operations. Marketing and Sales should then capitalize on this reduced production lead time to get higher prices and expand the market almost at will. It means Marketing and Sales have to be fully updated with the new capabilities and current status in production.

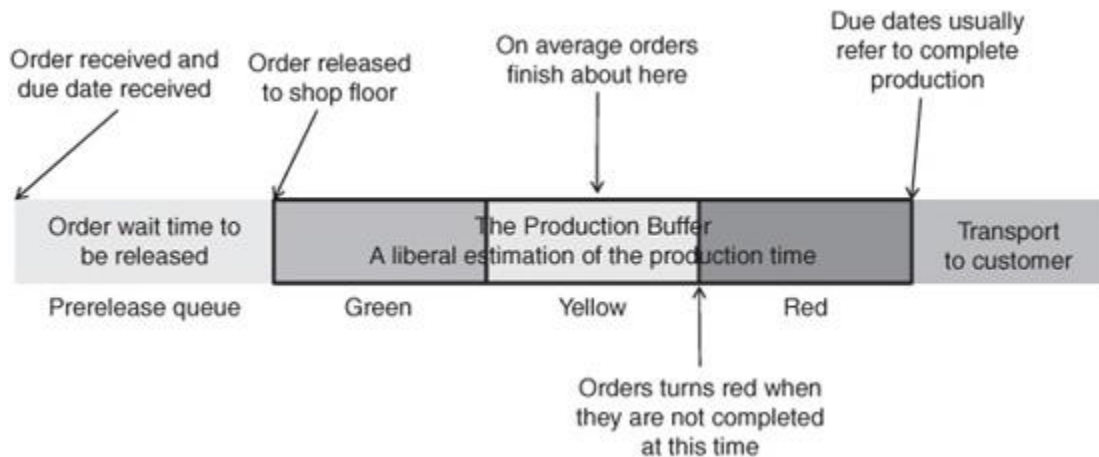


FIGURE 9-1 The elements of lead time.

Load Control

In traditional DBR, the role of the drum, the detailed schedule of the CCR, was to measure the load on the CCR and determine whether the promised due date was safe. When no detailed schedule is going to be used, then a replacement tool for measuring the load is required.

The *planned load is the accumulation of the derived load on the CCR, or any other relatively loaded resource, of all the firm orders (released and awaiting release) that have to be delivered within a certain horizon of time.* It is clear that more than one horizon of time might be defined. Each horizon provides a planned load for a specific use. However, a horizon is required for ensuring we do not promise delivery dates that cannot be met is of special importance. The time horizon for such a decision has to include all the orders already received (both released and not released to production) that might compete for the capacity required for the new received order. The important parameter is the realistic expectation of the market regarding the due date. If we assume that a client submitting an order expects to get it no later than the standard lead time of that industry, then

the horizon must include all orders to be delivered within the standard lead time. We might need to extend that horizon just a little to cover for a peak load that would force quoting a somewhat later due date.

There is definitely no wish for extending the horizon significantly more than the clients' expectations for response time to an order. It could be that some orders in the logbook have dates further out into the future, due to some clients wishing to ensure the availability of capacity for their orders. Nevertheless, do we need to consider those orders when we check the feasibility of delivering an order just received? Only when such an order enters the planning horizon should the capacity required for that order be considered.

What benefit does the planned load give? The most important information to deduce from the planned load is the approximate prediction regarding the time a new order will be processed by the CCR (the weakest link). Such information is critical for judging a "safe date" for completion of the order. It is only a gross approximation of the time the CCR would really process the order because our data are not necessarily precise and we do not guarantee the sequence of processing at the CCR. In addition, in order to obtain the planned load we simply add the load on the CCR of every order to be delivered in the horizon. Therefore, we have actually assumed that the CCR would not have idle time. See [Fig. 9-2](#). Thus, the timing when a new order would have the chance to be processed by the CCR is far from being accurate, but it can serve as a gross assessment. All we need to know is what due date is safe enough to promise completion of the order. For that, we need the planned load and to add to it a certain time buffer, as we'll soon see.

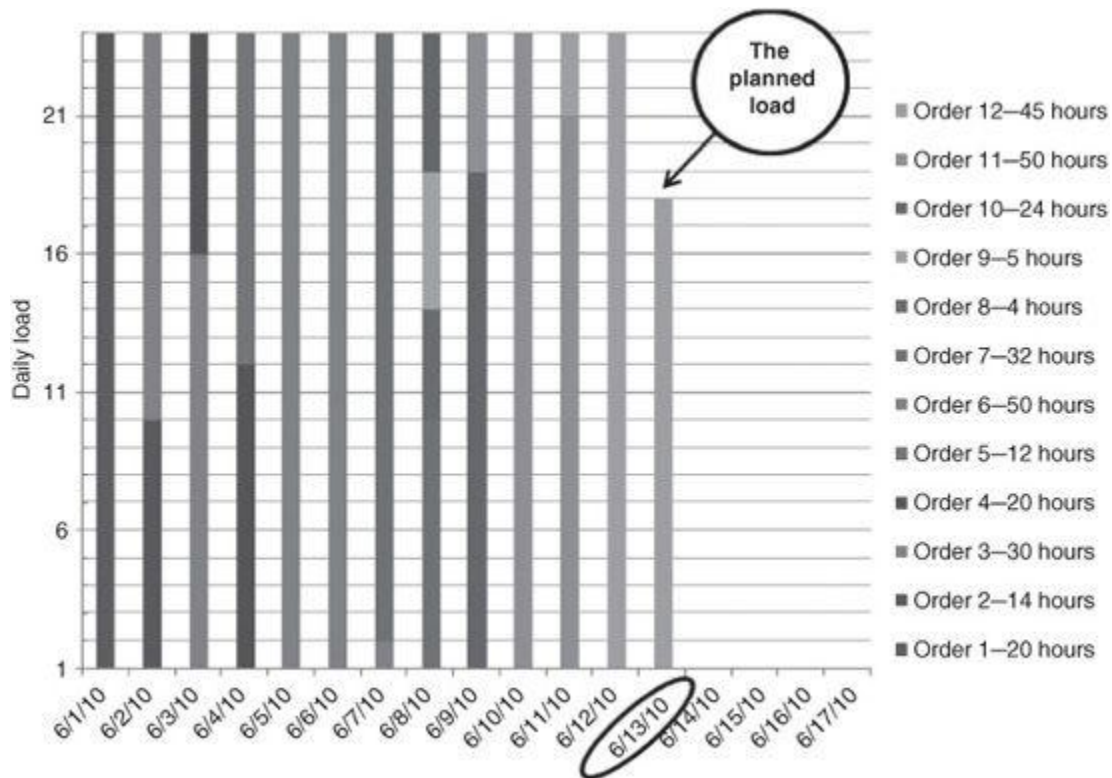


FIGURE 9-2 Demonstrating the calculation of the planned load.

Technically, the planned load looks like a schedule. It is generated by taking all the planned orders to be delivered and adding the required time the CCR has to invest on the timeline. The end of the planned load is a date—the approximate date when the CCR would finish processing all the currently known orders. In [Fig. 9-2](#), this date is 06/13/10. The important aspect of this date is that this almost arbitrary sequence is not forced on the CCR. The CCR is expected to follow the general priorities of BM and when some customer orders are stuck upstream, then other orders are to be processed early and the delayed customer orders would get higher priority when they show up at the CCR.

To demonstrate further that the planned load is not a schedule, let's review the following example. Suppose that

we are in a re-entrant environment where the order goes through the same resources several times. In the planned load, all the accumulated time the CCR has to invest in that order appears once. This view definitely does not provide a realistic schedule as the order is processed by the CCR several times and in-between these separate processing times other resources would be used to process this order. Suppose there are four machines, $M1 \rightarrow M2 \rightarrow M3 \rightarrow M4$, and a typical order goes through this sequence of resources six times. No matter which one of the four machines is the weakest link, there will be six times where the CCR would work on the order. Therefore, when such an order is put on the timeline of the CCR, it is placed on one continuous location where all the required capacity for six operations is included. On one hand, such a description is not realistic. However, as a gross average, this approximation is good enough for the time frame when the CCR would process the order and the total amount of capacity required. The first operation would probably be completed soon after the order is released to the floor, while the last one would be done significantly later. However, on average, all the operations will require approximately the time allotted in the planned load. We want to estimate the possible due date promised to the next order received. Then, the average time frame when the CCR would process the order is good enough, assuming, of course, the production buffer is long enough for the six iterations through the CCR.

In what sense does the planned load serve as a load control? The planned load represents a time frame that the CCR has to process every order to be delivered within the horizon. Therefore, the planned load, the date when it will finish on the CCR, should definitely be earlier than the horizon at hand. It needs to be shorter than the horizon by at least the time that is required by an order to move through the CCR to completion. For example, if the planned load finishing date on the CCR is three weeks from now and

the horizon is four weeks, then the minimal request for load control is that one week is more than enough for an order to flow from the CCR until completion of production for the order.

Surely, the planned load does not ensure that every single order, which is included somewhere in the order logbook, has enough time to complete production on time.

Determining the Safe Dates

Traditional DBR, like most production planning methods, assumes that once an order is received for production planning it already includes a mandatory completion date. Of course, there are times where a time quotation is required, but in most cases production planning gets the orders with the dates and then they must do their best to meet all of the requests on time.

In DBR methodology, there is a check, which is done immediately after the finite-capacity schedule of the CCR is completed, to validate that all the promised deliveries are secure. The check compares the time the CCR is scheduled to finish processing the order and the completion time for producing the order. The time difference must be equal to or longer than half of the shipping buffer. As a reminder, the shipping buffer in DBR is the liberal estimation of the required time between the CCR and order completion. Eventually, in the process of scheduling the CCR, that time difference cannot always be maintained. Sometimes the schedule provides more time than required for the order to pass from the CCR to completion, but many times less time than the shipping buffer is provided. Does it mean that when the time difference between the CCR schedule and the due date is somewhat less than the ideal shipping buffer the order is doomed to be late? The assumption used in that check is that as long as half of the shipping buffer time is

provided, then the completion time is secure enough because BM would give the order enough priority to pull the order through the remaining work centers to completion.¹⁵

In S-DBR, we need to develop the procedure for ensuring that the completion time can be secured. However, let's now challenge the assumption that production planning is given a due date and then, in some cases, Sales needs to contact the client and renegotiate that date. In other words, we must *let Sales know at all times what dates can be promised to any new orders coming in*. DBR production planning could not give a due date to an order that is not yet included in the schedule. Of course, one could always look at the DBR schedule and based on it assume what a safe date might be. Only later will the DBR schedule confirm the date or advise to delay the promised delivery date. In S-DBR, we can be much more flexible.

A procedure that would make it clear what dates can be safely promised to any incoming order generates important benefits. First, once Sales is convinced that those dates are not manipulated by Production, then it settles one of the main causes for the inherent tension between Sales and Production. Second, it opens a way to draw more of the CCR at times of peak demand by constantly smoothing the load on the CCR by giving dates that are based on the current load.

The rule for determining a safe due date is the *current planned load date (the first opening in the planned load)* plus half of the production buffer time.

In [Fig. 9-3](#) we see a graphic illustration of the time segments involved in computing both the safe date for committing orders to customers and for determining the release date for orders to production. Looking at the top right of the figure, we see that a full production buffer is placed at the safe date committed to the customer. Continuing at the top, we notice that we back off one full

production buffer length in time to determine the release date for the order. In the lower segments of the figure, we deal with the elements that go into the computation of the safe date. (Determine the date where the planned load ends on the CCR and add half of the production buffer.) We see that the production buffer (a liberal estimation of the production time) is logically divided in half. It can be seen that the point in time when processing is to occur on the CCR (depicted using a minimized picture of the CCR planned load profile) is the dividing point (in time) for splitting the production buffer in half. One half of the production buffer time is added to the planned completion date for the last order loaded in the CCR planned load, approximating when this next order will start processing on the CCR. This effectively adds the required time for completion on the CCR and mainly the end of the production process for the new order. Adding this half of the production buffer to the planned CCR date then gives us the safe date for commitment of the new order to the shipping dock. From that same point on the load chart, the other half of the production buffer (in effect the upstream half) is subtracted from the time when processing is to occur on the CCR. It is worth noting that in these calculations, the actual processing time on the CCR (touch time) is ignored in computing the release of raw materials and estimated shipping dates because CCR touch time versus queue time in the production process is usually negligible.

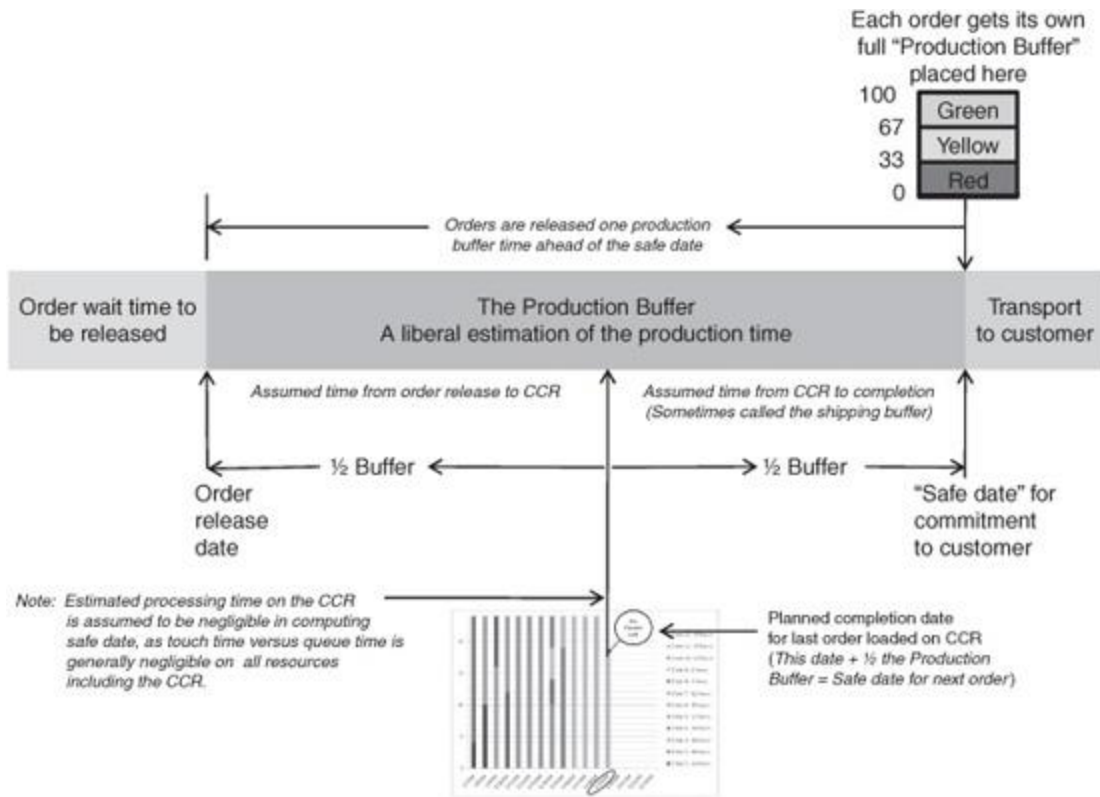


FIGURE 9-3 The timelines for safe date, order release, and buffer placement in S-DBR.

Taking this in another sequence, we see that from the point of order release to processing on the CCR we provide half the production buffer to get the order to the CCR, and the other half of the production buffer to get the order from the CCR to order completion. (Both of these time components are implied in releasing the order one full production buffer time ahead of the safe date.) Taking the time of processing on the CCR plus half the production buffer, we get the safe date for customer commitment.

The rule does not specify the location of the CCR operations within the routing. So, fixing the required time between the planned load and the safe completion date to be half the production buffer for the product requested by the order seems arbitrary. This works well except where the CCR is very close (measured in time) to the material release

or shipping in the product structure and in such an extreme case, some other division (of the production buffer other than half before and after the CCR) should be used as will be noted later.

Another question is whether the size of the manufacturing order has a major impact on the size of the production buffer and on the planned load. Suppose the production buffer for a “normal” order size of 50 units is 10 days. If the order is for 200 units, then it seems the processing time at the CCR would take significantly longer than usual, and it would impact the time the order requires from the downstream operations. Wouldn't it?

The rationale behind the rule of using the default buffer for orders with different quantities is to have a simple and straightforward method to monitor the load on the CCR and determine good safe dates. In a world where the data is frequently not accurate (processing times being, many times, just gross assessments) and being exposed to significant uncertainty from both external and internal sources, the only way to manage well is to look for “good enough” planning and execution rules. The time buffers introduced are usually not optimal, but as long as they are good enough, they do the job. A large order takes more processing time, but that time is usually small relative to the wait time, so as long as we do not speak of an ultra-large order, then probably the same production buffer should be used. In a case where an order, whose size is four times the regular size, is dealt with, then it seems reasonable that such an order has a higher chance of penetrating the red. If this occurrence is persistent, then we might introduce an increase to the production buffer when the specific order size is that much higher than the average order size. The same goes for considering the planned load. As it is not a real schedule and there is no guarantee that the CCR will process the order at the “scheduled” time, it is enough to assert that the given safe date is good enough and let BM

take the lead by establishing the required priorities. What else do we need to ensure? We need to release the order at such time that it'll have a full production buffer time to go through the whole shop, including the CCR.

Thus, the release time of the order should be according to the planned load minus half of the production buffer for that type of order.¹⁶ When the safe dates are given to the clients, then the order gets a full production buffer time to cover operations from the material release until order completion.

This rule applies for the vast majority of the cases. Of course, when the CCR is located at the very end of the routing (a very rare case¹⁷ in reality), then we adjust the material release and shipping buffer points by shifting the production buffer upstream by adding to the planned load end point just 20 percent of the production buffer to determine the shipping point. The same adjustment is used when a CCR is at the start of the routing. We release the material at least 20 percent of the production time buffer ahead of the planned load, and add 80 percent of the production time buffer to the planned load to determine the safe date for shipment. Only these extreme cases warrant deviation from the rule.

What Happens When Sales Quotes a Different Due Date Than the Safe Date Given by the Planned Load?

The recommended action is for Sales to quote the standard lead time whenever the safe date is prior to the standard. The problem exists when Sales does not follow the safe-date directive and quotes an earlier date than the safe date. If this case is a rare occurrence and most orders are quoted at safe date or later, then our recommendation is to release the manufacturing order to production one buffer time (production buffer) before the due date. If this practice of quoting due dates earlier than the safe dates is not a rare case, then it is not possible to use the safe-date mechanism

at all and the company must revert to the behavior of “we do our best to meet the dates, but sometimes we simply are not able to.”

Suppose the safe date is much earlier than the actual shipping date given to the client. Should the release date still be planned load minus half the production buffer? In this case, the actual time buffer is much longer than the production buffer.

The first impulse is to release the order one production buffer time prior to the committed due date. There is, however, one important reason why we should keep the release of the materials at the date the CCR is supposed to process it minus half the production buffer. If we do not release the order at that time, because it is more than the production buffer time until the due date, there is high probability that the CCR would be idle at that time or a little after that. One might argue that having the CCR idle, when it is not an active constraint is not as bad as it might seem. The fact that the date of the current planned load plus half the production buffer is earlier than the standard lead time date means sales are somewhat lower than what we are capable of handling. However, letting the CCR be idle when we do have a firm order at hand means we might lose that capacity if we'd need it in the near future. Let's view an example.

The standard lead time is eight weeks. The production buffer is four weeks. The current planned load is only two weeks. The rule says that the safe date is two weeks (the current planned load) plus half of four weeks (one-half the production buffer). It means the safe date is four weeks from now and the release date is now. However, if Sales quotes eight weeks, the pragmatic question is, “Should we release the order now, or wait for four weeks and then release it (leaving four weeks of production buffer for the order)?”

If we delay the release of the current order and do so to every order that gets the due date of eight weeks from now, then in two weeks the CCR would be idle for a while. If, within two to four weeks, many more orders would come, orders that would push the planned load to be more than six weeks, then due to lack of capacity at that time the quoted time for some orders might be more than eight weeks. Here is the damage—we do not fully utilize the CCR at the off-peak time and later we find we need the lost capacity.

Thus, the recommendation is to release the material based on the planned load minus half the production time buffer even for orders whose due dates are later than the safe dates. This ensures that as long as we have orders to be delivered within the standard lead time, we'll release them at the appropriate time to load the CCR continuously. Of course, there are obvious negatives to starting production earlier than the time truly required for safe on-time completion of that particular order. Nevertheless, given the situation that the shop floor is not fully loaded and wasting the capacity of the CCR might cause significant future damage, we still suggest releasing the order early in these circumstances.

Capacity Reservation

What If Some of the End Products Have Significantly Different Standard Lead Times?

As long as the planned load plus half of the production buffer is still at, or earlier than, the shortest standard lead time there is no problem—it is possible to quote the standard lead time. However, if the planned load stretches beyond the shortest standard lead time, what should we do? It could be the case that the short delivery can be promised because a substantial part of the orders have longer lead

times, so just a little manipulation of priorities, which is naturally done by BM (short buffer orders change their color faster), would still ensure excellent due date performance. But, how can we tell?

The problem in determining a safe date based on the planned load is that we assume that at the time we calculate the safe date all orders that are supposed to be delivered before that date are known. If, when we determine a safe date, we are not certain whether we have all the orders to be delivered until that date, then we cannot rely on the planned load.

There are four cases where this possibility of having shorter-delivery orders is relevant.

1. The perception¹⁸ in the market is that for standard products the delivery time should be shorter than for products that are more complicated.
2. A strategic client requires faster delivery than the standard.
3. The “rapid response” type of service, meaning clients are given an optional service of very fast delivery for a considerable markup.
4. Items that are MTS. The problem here is that when orders to stock are released, the time those items will be needed is not known a priori. We’ll deal with that case in another chapter dedicated to MTA.

All four cases have to be handled through the mechanism of reserving capacity for the “special” products that have to be completed faster. Suppose that on average the part of the “short-delivery items” requires about 25 percent of the capacity of the CCR. If we decide to assign only 70 percent of the CCR available capacity for the regular orders, so the

planned load calculations consider only 70 percent of the available capacity of the CCR, then we actually delay the calculated safe date for the regular orders.

Let's view an example. The CCR has an available capacity of 16 hours a day. The short-delivery orders take, on average, 25 percent of the CCR's capacity. However, as this is just an average, let's dedicate 30 percent of the CCR capacity to those orders, leaving only 11.2 hours a day (70 percent of 16 hours) for processing the "regular" orders. The calculation of the main planned load, used to determine safe dates, for regular orders should be based only on the regular orders and according to capacity availability of 11.2 hours a day. If the current regular orders contain 100 hours of CCR work, plus 27 hours for "special" orders, then the next regular order received should get a safe date of 9 work days ($100 \text{ hours of load} / 11.2 \text{ hours per day} = 8.92 \text{ days}$ rounded to 9 days) from now plus half of the production buffer expressed in work days. Of course, instead of the number of workdays from now, we should convert this to a specific date according to the calendar. This date is the earliest that could be given to the client as the safe date for completing the order.

What If a "Special Order" Is Received, What Safe Date Should It Get?

We assume the due date for a special order is not negotiable. It is solely dictated by the commitment to the market. This means that if the demand for special orders will grow beyond the reservation level, then the buffers of both the regular orders and the special ones would be consumed, which will cause a threat to the combined due date performance. In such a case adding capacity is definitely called for.

Should the Special Orders Get Higher Priority Than the Regular Ones?

The simple answer is no, all the orders compete on the capacity of the resources according to their color state (green, yellow, and red). If one wishes to be certain the special orders are on time, then increasing the production buffer for the special orders could be the right action.

An important point needs clarification: the CCR does *not* divide its time between the regulars and the specials according to the reservation percentage. The CCR works at all times based on BM priorities. When there are no special orders, then all 16 hours are dedicated to process regular orders.

Buffer Management

The basics of BM for an MTO environment did not change with the move from DBR to S-DBR. However, the criticality of BM to the success of the implementation has been increased. Once an order is released, then the only control on that order is through BM. The additional flexibility given to the execution phase makes providing good priorities an absolute must for successfully following the planning directives.

The planning procedure of S-DBR is that every order is given a production time buffer and a due date. According to these due dates, the material release schedule (due date minus production time buffer) is determined. At any given point in time, the time left for the order until its due date constitutes the remaining time buffer. The buffer time minus the remaining buffer is the portion of the buffer consumed thus far. The percentage of the buffer consumed to the total buffer is the "buffer status." Recall: Buffer status of less than 33 percent is considered green. Buffer status between

33 and 67 percent is yellow. Above 67 percent the buffer status is red.

An Example

The operator of a resource (not necessarily the CCR) has four different orders right now at the site. There are more orders in the shop, but only those four are currently at the work center. Order A1 should be delivered in 5 days; the production time buffer is 8 days. Order B1 is to be delivered in 3 days and the production time buffer is 6 days. Order C1 due date is 10 days from now and the production time buffer is 8 days. Order D1 has to be completed in 8 days from now and the production time buffer (which is long due to a lot of manual processing, fortunately most of the manual work has been done already) is 35 work days.

The color code for order A1 is

$$(8 - 5)/8 \times 100\% = 37.5\% \rightarrow \text{Yellow}$$

The color code for order B1 is

$$(6 - 3)/6 \times 100\% = 50\% \rightarrow \text{Yellow}$$

The color code for order C1 is

$$(8 - 10)/8 \times 100\% = -25\% \rightarrow \text{Green (actually above the green)}$$

Note that this order may have been released early to keep the CCR from going idle.

The color code for order D1 is

$$(35 - 8)/35 \times 100\% = 77.14\% \rightarrow \text{RED!}$$

Certainly, the resource should immediately process the D1 order. What will be the next one? We don't know at this time. More orders could show up and one or two of them could be red. Buffers and buffer penetration percentages for each of these orders are seen in [Fig. 9-4](#).

Several points regarding this example are worth discussion. Order D1 has the farthest due date from the current four candidates for immediate processing. However, because D1 has that long buffer, the assumption is that it needs that length of time as a buffer and thus the remaining 8 days signal the order is under pressure.

However, it is clearly written that the manual work in processing D1, which justified the especially long-time buffer, has already been dealt with. In such a situation, should we still look on order D1 as a "red order?"

It could be that at that point in the routing D1 is not truly urgent. However, we put a process in place that should yield good results in the vast majority of the cases. Can we really optimize the priority procedure to such a degree that in spite of the fluctuations in the shop and mistakes people make that better results would be achieved? Trying too hard to optimize within the "noise" (level of uncertainty) of the environment will actually increase the impact of the noise. This is one of the insights we have learned from Deming's funnel experiment.¹⁹

Suppose that D1 is stuck upstream of the relevant resource. Therefore, only three orders are at the site at this time. Which order should the operator choose? The buffer status of B1 is higher than A1, but the rule is to decide based on the color. The color of both orders is the same, yellow, and thus any choice of one of the two is acceptable. We assume the operator is aware of the buffer status, but might consider saving a setup, which is relevant only for orders with the same color code.

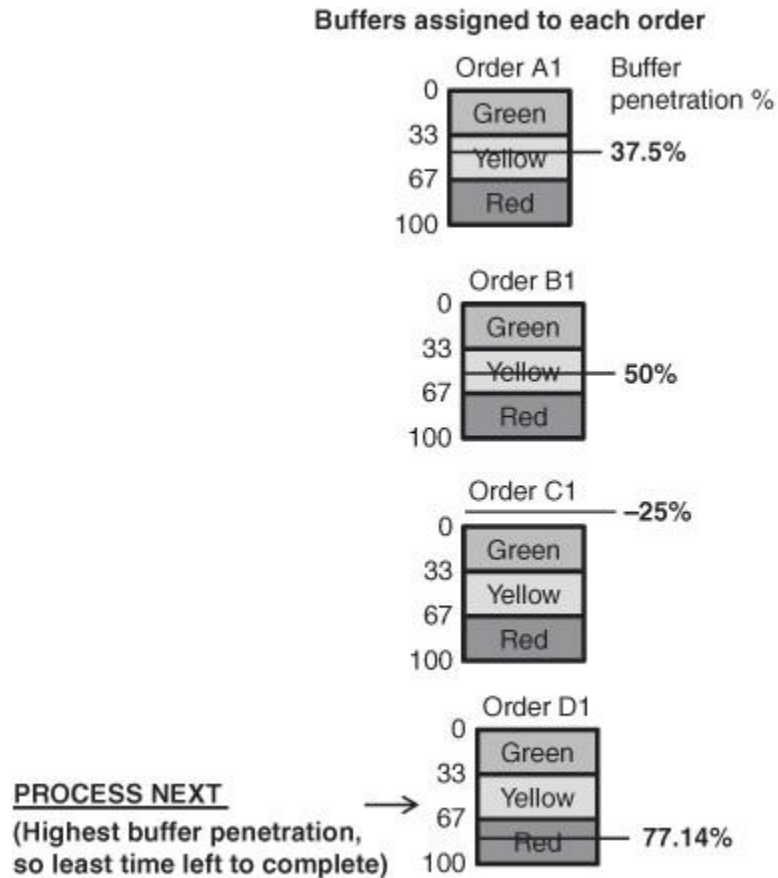


FIGURE 9-4 Buffer penetration shows priorities for processing.

Short-Term Planned Load

The main need for load control is for establishing the synchronization between Production and Sales, mainly by providing the estimates for safe dates as the earliest dates to be used by Sales for quotation. Thus, the horizon for that planned load is the expectation of the clients for response time. The planned load for this main horizon also provides signals for when to press Sales to bring more sales or restrain the sales to a degree.

What about the more immediate horizon, like the production buffer time? Every order that is now on the shop

floor has to be delivered within the production buffer time (unless that order was released early to keep the CCR busy). The point of load control at this time is to ascertain that the capacity of the CCR is more than enough to deliver everything on time.

How could it be that S-DBR would find itself in a situation where it ran out of capacity and there is not much hope to deliver all the orders on time? After all, every single order was given a due date that was in line with the available capacity as assessed by the planned load.

There are several possible causes for such a case of capacity shortage in the very short-term. One is that enough capacity is available, but it is in the shape of overtime or outsourcing, which means that capacity costs extra money. That extra capacity was probably considered by the planned load (depending on how the planned load was modeled), but now a clear decision is required whether to actually utilize the overtime and how much of it. Therefore, a short-term load control mechanism must be in place to support the decision on using overtime and how much overtime has to be used.

There could also be other, more severe, causes for short-term lack of capacity. It could be that Murphy caused downtime of the CCR, and that lost capacity is now taking its toll. Another cause is that too many special orders, to be delivered in a very short time, have been received and now it seems one or more orders would be late unless a quick way to add capacity is found.

A planned load for the short horizon, which is targeted at checking the capacity requirements for the short-term, should include all the orders released to the floor; that is, the regular and the special orders together.

Other special uses of the planned load include checking the special orders, taking into account only the reserved part of the capacity. That type of partial planned load is targeted to check the validity of the reservation level and

note the cases where the special orders would definitely “steal” capacity from the regular orders. Even if that situation is not problematic, like when the number of regular orders is not too high, the fact that the special orders have to steal capacity from the regular orders is meaningful enough to rethink the appropriate level of capacity reservation.

The Notion of “Slack”

As previously said, when the due date is set after the safe date given by the planned load plus half the production buffer, the order should still be released at the planned load minus half the production buffer. This creates an order with a larger production buffer than the regular one.

When the Sales regular policy is to quote the standard lead time unless the safe date is later than that date, then most of the orders would have larger time buffers. The software called *Symphony*, developed by Inherent Simplicity Inc., calls the time difference between the actual time buffer and the regular production buffer “slack.” Having “slack” means that Production is capable of processing additional orders while still meeting the current due dates. Thus, the slack is a signal to Sales to push for more orders.

There is another use for *slack*. What if an order is received with the request to deliver it sooner than the safe date? Normally we would expect that such a case is handled by the capacity reservation, but sometimes no capacity reservation has been accounted for, or the capacity reservation is fully utilized. However, if there are several orders with slack, then maybe they can be placed a little later on the timeline and thus provide an opportunity to deliver the new order at the requested time. What such software can do is simulate the updated planned load by inserting the capacity required for the new order in a place

that would support the requested date, and thus pushing some of the orders later in time, checking that all the orders still have the appropriate half-buffer between their assumed schedule in the planned load and their due dates.

Where S-DBR Fits Nicely

The original idea of applying S-DBR instead of DBR was that it fit the simpler environments. Certainly, it fits the case where no active capacity constraint is involved. As time went on, the understanding expanded to include within S-DBR cases where an active CCR existed but no sophisticated scheduling of the CCR was required. The assumption at that time was that when the detailed schedule of the CCR was straightforward, then it was enough to sequence the CCR according to BM priorities on the shop floor, but the more complicated cases should be handled by DBR and its three buffers. A good example of such an intricate case is one CCR operation feeding another CCR operation. The relatively complicated procedure for it is described in *The Haystack Syndrome* (Goldratt, 1990a).

The paradigm the author of this chapter had to fight with was that when the environment is truly complex, then detailed planning is a must. After all, there are many variables to take into account, so in order to achieve the required synchronization sophisticated planning has to take place.

Frequently inertia prevents one from breaking a paradigm. It was a startling experience to realize that such a common sense paradigm that a complex situation requires sophisticated planning has to be reversed!

Two meaningful citations of Goldratt, said at different times, have contributed to the change of paradigm:

1. In reality, we have both complexity and uncertainty and we are fortunate for having them both together.
2. The more complex the problem is, the simpler the solution should be.

The first saying means that when uncertainty is added upon complexity, then it is not possible to come up with an optimal solution that is also practical. When too many variables impact the output, then an “optimal solution” is usually a “sensitive solution,” meaning even a small deviation from the precise optimal solution would lead to a significant drop in the output. In an uncertain environment, there is no way to implement a multivariable solution without any deviation.

The following example should demonstrate the flaw in looking for the “optimal solution” to a complex and uncertain situation.

Suppose you live in Tel Aviv, Israel, and you need to arrive in Phoenix, Arizona, for a meeting with the board of directors of an important potential client. You would like to leave home as late as possible and spend the minimum time waiting for your connecting flights.

As the story goes, your agent has found exactly what you have asked for. You are booked on three legs of flights with 30 minutes between landing and takeoff. This time should be just enough to walk to the gate where the next flight is taking off. The agent included detailed calculations of the distance between the gates and the security and passport control in-between to show that it is an exact match to your capability to walk with your carry-on luggage. Eventually you are going to land at Phoenix Airport, 72 minutes before the meeting. The planning takes into account the traffic on the hour following the landing, showing that you will arrive at the meeting on the minute.

What do you think of such optimal planning? Isn't it ideal? It is great to be just on time without wasting precious time.

Even if we ignore most of the uncertainty and assume all flight schedules are truly precise, we need a lot of luck to arrive on such a nice optimal plan. Usually, the airlines and the availability of seats are not so generous and thus you'd need to waste some time waiting for flights or arrive significantly earlier than in the imaginary ideal solution. Yet, there are many alternatives to connect between Tel Aviv and Phoenix, so your agent should look for all of them until the best one is chosen.

Now, let's acknowledge the uncertainty. Flights are not always on time, queues for security and immigration fluctuate widely, and you can never rely on the traffic flow. At the end of the day, you realize not only do you need to consider buffering your plan, but also that there is no sense in checking too many options because once buffers are included in the planning the difference between alternative routings is negligible.

This is the insight emerging from the first citation of Goldratt: The inclusion of uncertainty within the complexity of the environment makes the sensible plan to be fairly simple. Focus only on the minimal requirements that are clearly critical and make sure those key requirements are properly protected.

This understanding leads to the second citation: solutions for complex situations must be simple; otherwise, they do not stand a chance in reality. Any small deviation in one of the many inputs (the number of inputs is what makes the environment complex) would cause too large a deviation in the output.

What does this have to do with S-DBR in a complex environment? The current understanding is that in the vast majority of the complicated cases, the use of S-DBR is even more sensible than the use of DBR. In the cases where there

is a problem in using S-DBR, the use of straightforward DBR is also problematic.

Take the case of a CCR operation feeding another CCR operation mentioned before as a re-entrant line. One has to assume a minimal time difference between the two CCR operations of the same order. This time difference is required to make sure the parts that finished processing by the previous CCR operation will reach the next CCR operation. In this way, several back-to-back time buffers have to be included in the planning, forcing long total lead time. When S-DBR is implemented in such an environment, there is no predetermined schedule for CCR. The practical consequence is that whenever the parts for the next operation reach the CCR they become available for immediate processing based on the priorities at that time. This allows the total length of the production buffer to be shorter than the total of the back-to-back buffers used in DBR.

Most cases that seem to require sophisticated scheduling of the CCR should use S-DBR as a practical approach that leaves most of the complexity to the last minute decision by the people who know the rules well enough and who are exposed to the real-time priorities as set by BM. However, S-DBR also has some limitations.

The Cases Where S-DBR Does Not Fit

S-DBR has two necessary conditions:

1. Arbitrary sequence of processing the orders does not significantly impact the capacity of the resources. In other words, the sequence as such does not cause any resource to become a bottleneck.

2. The ratio of the touch time to the production lead time is very small (less than 10 percent before S-DBR is implemented, less than 20 percent with S-DBR on). Touch time means the net processing time along the longest chain of operations. This definition is intended to exclude cases where assembly of thousands of parts, done on different sets of resources, might have a long processing time, but as the majority of the parts are assembled in parallel, the actual production lead time is not so long.
-

The environment where the first condition does not apply is where the length of the setup depends not only on what is going to be produced, but also on what has just been produced. Such a situation is usually called a *sequence-dependent-setup*.²⁰

When the difference between the setup times is very large, then S-DBR implementation might be problematic because an arbitrary sequence of processing orders, as dictated by BM, could easily turn a non-constraint into a bottleneck. The situation forces the producer to follow a certain sequence through the various products. Assuming that going through the whole cycle of products is a very significant portion of the standard lead time, the unavoidable result is that lead times might be quite long. What is even worse is that there is not much practical possibility to expedite any order because the sequence should not be changed. In other words, BM is able to show priorities but it is very difficult to follow them.

From this, one can deduce that even in a sequence-dependent-setup environment S-DBR can be applied if the total cycle time, the time between producing a product until it is possible to produce that product again, is short relative to the standard lead time. Another case where S-DBR is still fully applicable even in a sequence-dependent-setup

environment is when there are several production lines, which provide good enough flexibility to expedite an order without wasting too much capacity.

The second condition mentioned above makes us aware that manufacturing environments with relatively long touch times (more than 10 percent of the lead time before S-DBR is implemented) might pose certain difficulties in applying either DBR or S-DBR. In some of the cases, what is described as a manufacturing environment is actually a multi-project environment, where each single order is actually a project. In such an environment, the planning has to include sequencing of all resources within each project in order to define clearly the longest chain; otherwise, the lead time might be significantly inflated. Schragenheim and Walsh (2002) discuss the differences between manufacturing (DBR) and multi-projects (multi-project critical chain) and the appropriate planning and control systems for each. Critical Chain Project Management (CCPM) is the preferred method where the touch time is greater than 10 percent of the production lead time.

When the basic routing of every order is relatively simple, like a sequence of operations without parallel legs, S-DBR is still a valid option. However, having one or several very long operations poses some problems to BM to reflect the true state of urgency of the order. Let's use an example to demonstrate the problem.

Suppose the appropriate production buffer of an order is three weeks. The last operation is a long test that takes a whole week to go through. That one week is a fixed length of time. The sum of the previous net processing times is short, taking at most eight hours. If the test determines a problem, it usually requires the replacement of a purchased component, which is done in minutes. Therefore, all in all the touch time is approximately 35 percent of the production buffer, but the vast majority of it is accumulated at the very end. What is the appropriate priority of an order

after two weeks? The regular priority would show an order just entered into the red. But, if the order testing has not started yet then actually the order is already late.²¹ On the other hand, if the order is already three days into the testing then it'd most certainly be on time.

In the previous example, only one operation is truly long, while the rest are normal manufacturing where the net processing time per piece (also per order) is very short. In this case, it is possible to introduce some changes to the way S-DBR and BM rules are implemented that would work. In the particular case of the example, all we need is to model the requirement to reach the testing no later than a week earlier than the customer's due date. By generating a fictional safe date not for the full completion of the order, but for entering the last operation one week before the due date, we force the right priorities in the system prior to the final testing.

In cases where the long operation is in the middle of the routing, there might be a need to implement back-to-back time buffers, creating a somewhat less simple solution, but still not requiring scheduling the CCR in detail, and most of the S-DBR procedures are in place.²²

In the extreme case (such as firms in the process industry) where an order has several long operations that are spread through the routing, yet the routing itself is a simple *'I'-shaped structure*,²³ changes to the buffer management algorithm would still provide the right priorities. This feature, developed by Inherent Simplicity Ltd., is beyond the scope of this chapter.

Implementation Issues and Processes

One of the primary advantages of S-DBR over the traditional DBR is in the speed of implementation and results.

Implementing S-DBR should always start with choking the material release so that only orders to be delivered in the horizon of the production time buffer would be found in the shop floor. We've already mentioned that a good initial estimation of the time buffer is one-half the current production lead time. If it is not clear what the current production lead time is, then take the standard lead time in this industry and cut it to half.

A few exceptions to the half-the-production-lead-time rule exist. The first is an environment of a dedicated assembly line, where all the WIP in the line is restricted to several hours. The other exception is where real effective Lean methods have vastly reduced the WIP and lead time. In those cases, the production buffer can be based on the current production lead time for implementation.

Choking the material release must include dealing with the current batching policies by either abolishing them by making the customer order quantity the batch size or at least reducing the batch size.

The next mandatory move is to establish BM. This move can be done manually or be supported by software. Putting red labels on the red orders is a simple visual tool for an initial implementation of BM. For the operators in the shop, the rules of behavior with red orders must be absolutely clear: The workers must take responsibility to flow the red orders to completion. If an operator needs materials, tools, drawings, or anything else required to move a red order, he or she must get whatever is needed or notify production management immediately of the support needed. Overtime is another option of production management to deal with red orders.

Implementing the load control function generally takes a little more time. Having the planned load in place is not mandatory for getting the initial results. Actually, it is not even very urgent to identify the "weakest link" (the relative CCR). The assumption is that the demand would not rise

very quickly, and thus choking the release based on the production buffers and the simple priority rules are enough to improve the due date performance and stabilize the shop.

Once the implementation stabilizes the shop, then identifying the CCR is easy enough. It is the resource that most of the time holds the longest queue of WIP as measured in processing time at that resource. Then the initial steps to implement planned load can be taken. If more than one natural candidate for the CCR shows up, then monitoring the load on three to five work centers is good enough. Once good data on the planned load is obtained, the identity of the real CCR becomes clear. If more than one CCR exists in the same flow, then determine logically which should be the one to use and increase the capacity of the other.

The next step is to establish the rules for Sales to quote due dates, which considers the safe dates given by the planned load (plus half the production buffer). Now the implementation is ready to face a real increase in sales.

The *process of ongoing improvement* (POOGI) should be established at this point. The idea is that every time an order becomes red, a reason should be entered by a person in charge from the production management personnel. The reason is taken from a prepared table of possible reasons. A reason must answer the question "What now delays the order?" The list of reasons (such as "Quality problems are identified and being taken care of," "Huge queue of work at work center X due to a long machine breakdown," "Work center X currently works on the order," etc.) is presented weekly as a Pareto list and a team under the direction of production management should look to eliminate the top causes of lateness on the list. This procedure should improve the flow even more and then efforts to capitalize on it by creating offers that are more lucrative to the market should be taken.

Looking Ahead to MTS

This chapter focuses on MTO environments. DBR, like previous production planning methods, has assumed every production order must have a due date and that due dates determine the relative priority of any production order. The next chapter is going to show that this assumption is not necessarily true and actually, there should be a clear distinction between MTO and MTS, where no definite customer order exists at the time of material release to production. The next chapter will also deal with mixed environments where certain products are MTS, while others are MTO.

Suggested Reading

Schragenheim, E., Dettmer, H. W., and Patterson, W. 2009. *Supply Chain at Warp Speed*. Boca Raton, FL: CRC Press. [Chapters 3](#) through 5 are especially relevant. www.inherentsimplicity.com/warp-speed is a site that allows downloading of the MICSS simulator including analysis files and more related materials.

References

- Blackstone, J. H. 2008. *The APICS Dictionary*. 12th ed. Alexandria, VA: APICS.
- Fry, T. D., Cox, J. F. and Blackstone, J. H. 1992. "An analysis and discussion of the OPT[®] software and its use," *Production and Operations Management Journal* 1(2)Spring: 229-242.

- Goldratt, E. M. 1990a. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Crotonon-Hudson, NY: North River Press.
- Goldratt, E. M. 1990b. *What Is This Thing Called Theory of Constraints and How Should It be Implemented?* Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 1997. *Critical Chain*. Great Barrington, MA: North River Press.
- Goldratt, E. M. 2009. "Standing on the Shoulders of Giants," *The Manufacturer*. June. http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants. (accessed February 4, 2010).
- Goldratt, E. M. and Cox, J. 1984. *The Goal*. Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. and Fox, R. E. 1986. *The Race*. Croton-on-Hudson, NY: North River Press.
- Schrageheim, E. and Dettmer, H. W. 2000. *Manufacturing at Warp Speed: Optimizing Supply Chain Financial Performance*. Boca Raton, FL: CRC Press.
- Schrageheim, E., Dettmer, H. W., and Patterson, W. 2009. *Supply Chain Management at Warp Speed*. Boca Raton, FL: CRC Press.
- Schrageheim, E. and Walsh, D. P. 2002. "The critical distinction between manufacturing and multi-projects," *The Performance Advantage*, February, pages 42-46.

Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>.

About the Author

In the last 25 years, Eli Schragenheim has taught, spoke at conferences, and consulted in more than 15 countries, including the United States, Canada, India, China, and Japan. He has also developed software simulation tools especially designed to experience the thinking of TOC, and consulted to several application software companies to develop the right TOC functionally in their own packages.

Mr. Schragenheim was a partner in the A.Y. Goldratt Institute and he is now a Director in The Goldratt Schools.

He is the author of *Management Dilemmas*. He collaborated with H. William Dettmer in writing *Manufacturing at Warp Speed*. He also collaborated with Carol A. Ptak on *ERP, Tools, Techniques, and Applications for Integrating the Supply Chain*, and with Dr. Goldratt and Carol A. Ptak on *Necessary But Not Sufficient*. In March 2009, a new book titled *Supply Chain Fulfillment at Warp Speed*, with H. William Dettmer and Wayne Patterson was published. The new book contains much of the new developments of TOC in operations.

Mr. Schragenheim holds an MBA from Tel Aviv University, Israel, and a BSc in Mathematics and Physics from the Hebrew University in Jerusalem. In-between his formal studies, he was a TV director for almost 10 years. He is a citizen of Israel.

CHAPTER 10

Managing Make-to-Stock and the Concept of Make-to-Availability

Eli Schragenheim

Introduction

Is there a basic difference between producing an order for a specific customer order and producing an order in anticipation of future demand? From a business perspective, there is an obvious difference: producing in anticipation of demand means risk, while producing to a firm order looks safe enough. However, once there is a decision to produce to stock, either based on formal forecasting or on a hunch, should there be a difference in the rules behind production planning and execution?

The traditional approach does not see much of a difference between *make-to-stock* (MTS) and *make-to-order* (MTO) for production management. Thus, mixing within the same work order a quantity that is covered by firm orders with a quantity based on anticipation is very common.

When *Drum-Buffer-Rope* (DBR) (Goldratt and Cox, 1984; Goldratt and Fox, 1986) was developed in the 1980s, it did not challenge the assumption that there is no difference in planning the shop floor for firm orders and planning it for anticipation of future demand. In addition, *Buffer Management* (BM) did not see any difference between MTO and MTS.

This chapter argues that there should be a difference. It is designed to explain the logic of why different rules, both in the planning and in the execution, are required and goes on to detail the method itself and its ramifications.

While dealing with the topic of MTS and how it is different from MTO, another insight by Dr. Goldratt has emerged that led to a new term called *make-to-availability* (MTA), where we add to the operational meaning of MTS a marketing message: *We commit to our chosen market to hold perfect availability of a group of specific end products at a specific warehouse.* The objective of MTA is to offer a new business opportunity based on providing extra value to clients through guaranteed lead times, which competitors will find it hard to imitate.

Copyright © 2010 by Eli Schragenheim.

In this chapter, we explain the operational ramifications required to offer this commitment to the market. We do not go into the marketing side¹ of how such an offer could be used to enhance the perception of value of the client and how to capitalize on that added value to gain more profits to the organization.

The chapter deals with why there is a need to change to the *Simplified Drum-Buffer-Rope* (S-DBR) methodology and the related BM mechanism, presented in [Chapter 9](#), to deal with MTA. Then we present the methodology itself, both the planning and the BM rules. Following that we deal with some broader issues of MTA, like managing seasonality and mixed environments of both MTA and MTO or cases that are MTS rather than MTA. Toward the end of the chapter, we highlight some practical implementation issues.

Why Is a Special Methodology for MTS Required?

Two different parameters are usually considered in evaluating planning the production of MTS. One is determining the quantity to be produced and the other is fixing the date for the shop floor to complete production.

Is anything a little bit strange in the second parameter?

When a client submits an order, the due date is important. Does the client truly need the order on that date? Moreover, even if the client needs some of the ordered quantity at the agreed upon delivery date, in most cases not all of the quantity is required at that time. Still, he has the right to expect delivery at the agreed upon date and missing that date can cause negative effects on the reputation of the supplier. Therefore, it is natural that efforts should be made to deliver all firm orders on time.

Is it really the same in producing to stock?

The required quantity in producing to stock is an estimate. The chosen quantity to produce is not likely, in most cases, to be consumed at the date given to the production order. Therefore, the date simply sets the priority for the order and the performance measurement for the shop floor.

Let's see if this date is good enough for setting the internal priority in the shop floor. What truly dictates the priority of a production order for stock? In most cases, production to stock aims to provide availability of the item to any urgent order. In such a case, the true priority of the production order has to be dependent on the availability of finished stock for that particular item. Will stock be there for the urgent order? In addition, if the due date performance for MTS orders is in most cases not very critical, should we make it the prime performance measurement?

Our conclusion is that for MTS there is a need to redefine the priorities for the shop floor and base the appropriate performance measurements on that. This means we might have to develop a different BM scheme for MTS.

There is one case though where an MTS order has to be completed at a certain meaningful date. This is when the

stock is to provide availability at a date where we anticipate a significant demand, like a holiday or the first day of an advertised promotion. In this type of MTS, the date is very important. However, in all other cases, certainly when the point is to support continuous availability of the items, the required date has no special meaning.

The decision on what quantity to produce to stock is also quite different from MTO. The *Theory of Constraints* (TOC) is focused on generating Throughput, which is not the same as generating output. Therefore, while in MTO the client's wishes, as expressed by the firm order, dictate both the quantity and the completion date and directly results in Throughput, for MTS we need another approach.

The Current Confusion in Managing Stock

The current practice in production management inter-mixes MTO and MTS. *Economic Order Quantities* (EOQs) lead production planning to fill the demand for current customer orders and then add stock intended to cover future orders. This combination of customer orders and stock orders executed in a *material requirements planning* (MRP)² environment uses the problematic notion of the "available to promise" algorithm. This algorithm helps in deciding whether current requests can be reasonably met in quantity and time. The problems with this algorithm³ are twofold: the first is the unreliable way uncertainty in the shop floor is handled, and the second is inconsistency due to the varying levels of stock that is not already assigned to firm orders. From the potential customer point of view, sometimes an order is delivered very soon and sometimes an order is delivered relatively slowly. This is problematic because there is no standard for the customer to rely on.

What makes the mix between orders and stock even more confusing is that in MRP every pass from one level to the

next level in the *bill of materials* (BOM) has its own work order, which often merges the requirements from several customer orders and then inflates the work order even more by adding items for stock. As the expected customer demand changes at the top level, those fluctuations are then exploded to the lower levels in the BOM structure with each new iteration of MRP (often done weekly) thus impacting the ratio between the parts that are required for firm orders and parts that are for stock. This means how much component stock for future parts has been added is arbitrary and not derived based on a calculated decision to maintain a certain level of stock of a specific component. In this way calculating the “available-to-promise,” looking for the available stock of a large number of components, is very tricky indeed. It could easily be that for a certain end product some of the required components have a lot of stock, while other components are short. MRP developers have tried to treat the effect of this nervousness by providing *pegging* (Blackstone, 2008, 97) “*to determine requirements traceability, which allows one to trace the source of requirements through record linkages.*” (© APICS 2008, used by permission, all rights reserved.)

Another source of confusion is the reliance on forecasting, or rather the common misunderstanding of how to use forecasts to support good decisions.

The Common Misunderstanding of Forecasts

The forecasting algorithm is not a prophecy and was never intended to answer questions like, “How many units will be sold next month?” Forecasting is a statistical model that describes, under certain assumptions, a specific uncertain future behavior of a specific variable. Being just a statistical model means all it can do is point to a possible spread of results treated in a solid statistical way—finding a probable

average and a probable standard deviation around that average. By providing this partial information on the possible range of results, it allows the decision maker to consider where in the range it is best to place the quantity in question for minimum risk.

The common misunderstanding of forecasting has two parts. The first is to understand what partial information the forecast should provide. The second is how to make a good decision based on the forecast information.

The common ignorance regarding the first part is focused on *using the forecast as a single number*. The mathematical/statistical handling of all uncertain functions includes, at the very least, two parameters. The common minimum description of uncertain behavior is the use of the *average* and the *standard deviation*. Another option is to describe a spread of possible results by *the confidence interval*: a range of results that encompasses, according to the forecasting assessment, 95 percent or more of the possible results.

The common use of the forecast as a single number is causing huge confusion because the essential range of results is missing. Thus, it is almost useless and definitely misleading. The vast majority of management reports contain only the column of the forecast, namely the average predicted forecast. The forecasting error, the equivalent of the standard deviation, is not mentioned in those reports. The basic misunderstanding is even more destructive when people, mainly from Sales, are required to give their "forecast" for the next period. What kind of single number does management hope to get? The average? Do they really get a fair assessment of the average from the salespeople? Could it be that a typical salesperson provides his intuition regarding what he hopes to sell, rather than his or her estimation of the average? The salesperson wants to be sure he would have available all the quantity he might sell. On the other hand, the salesperson may want to give his

estimation on what he is sure he can sell and not be caught failing to meet his forecast. The point is that for an unclear question, people get answers to whatever interpretation the person answering the question has in mind.

Is the “average” forecast the required information for a decision regarding how much to produce for stock? Let’s consider the following example.

The forecast for next month sales is 1000 units. We should have in stock, at the start of next month, 300 units (also “on average” depending on the sales until then). Assuming the policy is to produce the whole monthly requirement in one batch (usually an unwise policy, but that is not the point right now), should we produce 700 units?

Well, if that is the only information we have, then we are led to make a faulty decision. A proper forecast should also contain, at least, an indication of the forecasting error. Suppose the forecasting error is 500 units. The hidden meaning is that it is perfectly possible that the real demand next month would be 1500 units. Even 2000 is still a valid possibility. Of course, it also means you might sell only 500 units. Managers are required to make sound decisions even when living in such an imperfect world. A much better decision than to produce 700 units would be to produce 1700 units to cover for the valid possibility of having demand for 2000 units. Another sound decision could be to produce only 200 units when the concern of being left with unsold products is more severe than being short of products. In other words, *any sound decision has to take into account the damage of producing too much versus the damage of producing too little* and the larger damage should dictate whether to produce more than the average or less than the average. In most cases, the decision to produce according to the “average” prediction (based on a single average forecast number) is a truly bad decision because the element of risk is not brought into the picture. Suppose that the plan is to produce more than the average.

However, without any indication of the possible spread, how should one make up his mind regarding how much more to produce?

Misunderstanding of the forecast has more aspects to it. Forecasting the sales of just one item in the coming month might be too “erratic”; thus, the idea is to forecast the sales of the whole product family. This should yield a much better forecast, shouldn’t it?

Well, usually the term “better forecast” means a relatively smaller forecasting error, while the term “erratic forecast” means a very large forecast error. The problem is you cannot use that “better forecast” for a better decision on the level of the individual item. Suppose it is “known” that the sales of a certain item are approximately 10 percent of the sales of the total family. Do we get a “better forecast” for that item when we take the forecast for the sales of the family as a whole and then take 10 percent of it as the forecast for that individual item? No, you do not get a better forecast for the individual item this way. You have a gross estimation of the average, but the possible spread of the results for different units in the product family is pretty high and you cannot reduce the spread of the sales of an individual item by forecasting the whole family. For a decision on the demand level of an item, one needs its forecast including the assessment of the spread of results.

Another aspect of ignorance regarding forecasting relates to the forecasting horizon and the time periods within that horizon. Management likes to look at “the big picture” and thus wants to see not just the forecast for next month, but also for the subsequent months—at least up to one year. Suppose the forecast for next month is 1000 units plus or minus 500 units. Now, if the forecast for the month after that is also “on average” 1000, the “plus or minus” is probably larger. The farther in time we go, the larger the spread of the forecast. There are two reasons for it:

1. Naturally when estimating, the forecasting error gets larger for subsequent periods because any deviation in the trend of the sales would get larger the farther out in time we look (increased uncertainty).
2. The most troubling point in forecasting is based on the assumption that the characteristics of the past are not going to change, and thus we can deduce the future from the past. As we look farther in time, there is a higher chance that an event will change the basic parameters. Just consider the case where today your main competitor has opened his manufacturing facility not far from you and he is going to try to move your clients to him. Suddenly, the rules of the game change and you cannot rely on the past to deduce what your future sales will be.

The direction of the solution to forecasting demand is hidden exactly in the notion that for the very short-term we have a good idea of what is going to happen. Even when we look at the short-term, we must consider not just the average, but also how much *we might sell*. In other words, we need a confidence interval to give us a reasonable range so we can decide how to prepare ourselves to a valid level of sales. When response time is rapid, there is no real need to forecast beyond the short-term except to look at approximations of capacity, materials, and cash requirements.

The Current Undesirable Effects in MTS

Every time a production order is released without a definite customer order, it might create a surplus of inventory and at the same time delay the production of another product, which might be in high demand on short notice. There is no

way to avoid these mistakes, simply because we are not prophets and we cannot really know the future.

The unavoidable result is that at any given time the finished-goods inventory of some of the products is excessively high relative to the actual demand, while there are shortages of other products. All we can hope for is to eliminate the shortages to such a low level that we have almost “no shortages,” while the surpluses of inventory are rather limited.

The current state causes many other undesirable effects within the shop. Holding too much stock takes its toll on financials, limits space for other items, and causes pressure to “get rid” of stock. Producing based on misunderstood long-term forecasts leads to producing very large batches (“we should produce this product only twice in the year to gain efficiency”), which causes long production lead times and delays for products that are truly required by the market.

Already mentioned is the confusion between MTO and MTS that leads to inability to state clearly when a current request can be met safely. Another common undesired effect happens at the level of common components that go to many end items. Components are often “stolen” by the overproduction of end items with low demand, while other items are short. What makes the “stealing” effect special is that it creates anger and tension because the cause and effect are visible to employees. One can clearly see how the decision to produce too large a batch has exhausted all the necessary components for a truly urgent order.

What to Do? The Direction of the Solution

The Basic Principle of Flow

The immediate conclusion in understanding the characteristics of forecasting should be: *the faster we can respond, the more reliable is the forecast*. Embedded in that sentence we have the recognition that we should not look too far into the future when framing production orders. However, there is a minimum time into the future where we need to ask the question: *how much might we sell?* The default assumption is that we want no shortages and for that, we are ready to pay the price of holding more inventory than we would need in a world with perfect knowledge. Therefore, our aim most of the time is to have full availability of those items on which we choose to maintain excellent availability, while the amount of stock is nicely controlled at a level that is still appropriate for preventing shortages.

The question points to the need for a different type of forecast, not the regular one. The question is not directed at the average sales within the response time, but at what we *might actually sell*. In other words, forecasting the maximum sales that we could reasonably expect for that period of time. In order to fully support availability while refraining from overproduction, two practical insights emerge:

1. Production still needs to focus on the flow to finished-goods inventory, flowing the required quantity as fast as possible through the shop.
2. Unless we have a good reason to believe that the market demand is going to change, or that the current inventory in the system is either too high or too low, then a simple straightforward reaction to any sale is to replenish that quantity. This means that replenishing the exact quantity of what was sold is a natural default. From the production planning perspective, it

means that everyday production should initiate producing the exact quantity of what was sold yesterday.

From the two insights, it is clear that we need to determine the appropriate inventory in the shop that would provide perfect availability, thus maintaining the fastest flow of goods to customers. The other critical point is to improve, and keep improving, the internal flow.

Another understanding emerges. If the objective is to provide perfect availability to customers, then we should state that openly and probably take one more step and commit to maintain that availability openly by letting our customers know that is our commitment.

From MTS to MTA

The verbalization of the ultimate objective is: *We commit to our chosen market to hold perfect availability of a group of specific end products at a specific warehouse.*

This objective has two critical elements in it. One is the marketing message, defining the target market, the items that it includes, and possibly also some limitation of the one-time demand that such an availability would provide. The other is the operational element. Once a commitment is given, Production must perform to meet that commitment.

Let's now clarify the relationship between MTS and MTA. Certainly any case of MTA requires MTS, unless production can be done in a few seconds. However, many cases of MTS are definitely not MTA. Those cases happen every time there is no concrete commitment to availability.

Let's present two different examples for MTS that is not MTA.

Example 1 Painters, including famous ones, paint regularly for stock, meaning without a specific client commissioning the painting. However, many paintings are

done only once, there is just one unique single copy. In some cases, a limited number of copies (authorized prints) are produced. This definitely is not a commitment for availability. Similarly, exclusive fashion items are also promised to be single units; there is no promise of availability.

Example 2 Items that are going to be sold in a specific period of a few days, for instance, souvenirs for a specific sporting event, like hats or T-shirts, with the appropriate logo and colors for one of the teams for the big final game, will need a lot of stock before the event. After the event, sales will be very low. The time those items are sold is so short that there is no practical chance to replenish. In such a case, there is no clear commitment for availability. Actually, the producer hopes to sell all his items (while not losing too many sales) and usually that means leaving some demand unsatisfied.

Determining the Appropriate Inventory

The idea of replenishing exactly what is sold, or more accurately what is consumed,⁴ has an interesting ramification: the inventory in the shop remains fixed. However, it is fixed through the whole shop floor, both at the finished-goods and work-in-process on the shop floor itself. Therefore, the set of parts and assemblies needed to complete the end product committed for availability may exist in various levels of completion on the shop floor. When fully fabricated, this inventory would equal the quantity needed for committed availability.

Of course, here and there the total inventory might be less than the regular fixed amount because Production is slow to release the next work orders, but the concept means trying to keep the total stock fixed.

There is a lot of sense in determining a fixed quantity per item to protect availability. Ideally, it'd be best to keep fixed stock in finished goods. However, this is quite impossible because once there is demand the stock goes down. Then what do you do? The only way to react to actual demand is to initiate replenishment. Thus, defining the "shipping buffer,"⁵ the protection mechanism that protects the availability, as the total amount of finished goods plus the *work-in-progress* (WIP) is the simple and straightforward way to institute the appropriate protection mechanism.

Let's call *the buffer, the fixed stock in the system, the target level* for this item. How should the target level be determined? From the time one piece is sold until the time the replenishment piece arrives at the finished-goods warehouse, availability has to be maintained. Let's describe *the average time it takes to replenish a piece* as the *replenishment time*⁶. Most certainly, the target level should include the average demand within the replenishment time, but this is definitely not enough, as there is a need to address what *might be sold* and also to address any case where the replenishment will take more than the average replenishment time.

There are two practical ways to handle the determination of the target levels. One way is to take the above average demand within replenishment time, which is information that is usually easy to get, and multiply it by a "paranoia factor" to include peaks of sales and certain blockages in production. In the area of the production floor, a minimum "paranoia factor" is an additional 50 percent (factor of 1.5 of the average). Using this number is recommended in situations where no sequence-dependent setups exist, thus managing the priorities on the floor (still to be discussed in this chapter) toward rapid work flow. In other cases where the demand fluctuations are especially high and blockages in the flow are frequent, a factor of 2 should be used.

Another approach looks at the recent 6 to 12 months of history for the actual maximum sales that have occurred within that window of time defined as the *reliable replenishment time*. The reliable replenishment time means *that when you really need it you can safely get it within that time.*⁷

Two important points to notice at all times:

1. If currently there is no stock, or very little, in finished goods (do not count stock that is already assigned to clients), then *first build the finished-goods stock* and only later move to the TOC MTA solution. More on that point later, but please take note!
2. The determination of target (maximum) level based on the appropriate criteria as discussed above is only to set the *initial* inventory levels. As we'll see, future changes to the target levels (increase or decrease) are done based on a special algorithm that monitors the actual behavior of the finished-goods stock.

Buffer Management in MTA

Once the target level is operational, daily replenishment orders are initiated based on the previous day's consumption. *Every production order for replenishment is released without any due date.* The immediate question is how should the priorities on the floor be determined? There is no doubt that there is a real need to settle the priorities.

The idea is that the priority of the orders depends on what lies downstream of the production order, in other words between the order and finished goods. The amount of stock not dedicated to any client, and therefore available to fill new customer orders relative to the buffer size (the target level), is the real indication to how urgent the production

order is. We do not really expect to have 100 percent of the target level completed and in the finished-goods warehouse. This would be way too much as we expect the replenishment to arrive at the warehouse much faster than the time the whole target level is going to be consumed.

Let's look at the situation shown in [Table 10-1](#). It shows the full picture of a target level inventory buffer for a product P1. Suppose that a production order for 200 units for product P1 lies somewhere on the shop floor. Downstream from that order is the finished-goods stock, which contains 100 units. Suppose that the target level, the amount of inventory we believe would provide excellent availability, is 500 units. We know that the whole target level should be in the production system somewhere, either in finished goods, or at some level of completion within the shop floor. This means that right now only 20 percent of the target level actually resides at the finished-goods inventory. It looks like replenishing the finished-goods stock is urgent. Note also the fact that the size of Order 1 of 200 units is not required for the assessment of how urgent the order is. The question of urgency relates to how much is in finished stock downstream from Order 1. Like BM in MTO, we like to denote the priority of any order by a color code: green, yellow, or red. Color code definitions are defined more fully next in this section. They are shown in [Table 10-1](#) to complete the picture in our example showing the relative priority of the orders. Order 1 is urgent and is in a red buffer status. Order 2 is upstream from Order 1. It has 300 units downstream from it, or 60 percent of the target inventory. It is in a yellow buffer status. Order 3 for another 100 units has 80 percent downstream in front of it, 80 percent of the target, and a buffer status of green.

Inventory and Production Orders	Quantity	Percentage of Target in Front of Order (Downstream)	Buffer Status (Priority)
Finished goods	100		
Order 1	200	20	Red
Order 2	100	60	Yellow
Order 3	100	80	Green
Target Level	500		

TABLE 10-1 Availability Targets and Priority Status of Orders for a Buffer Target of 500

Defining Buffer Status

We define the state of the finished-goods buffer when containing two-thirds or more of the target level as green. In other words, one-third or less of the buffer is not in the finished-goods inventory, but somewhere on the way.

In a similar fashion, when the finished-goods inventory contains between one-third and two-thirds of the target level, as shown in [Fig. 10-1](#), we call that state yellow.

When the on-hand stock, the inventory at the finished-goods warehouse, is less than one-third, meaning more than two-thirds are not at the warehouse, then the state is red.

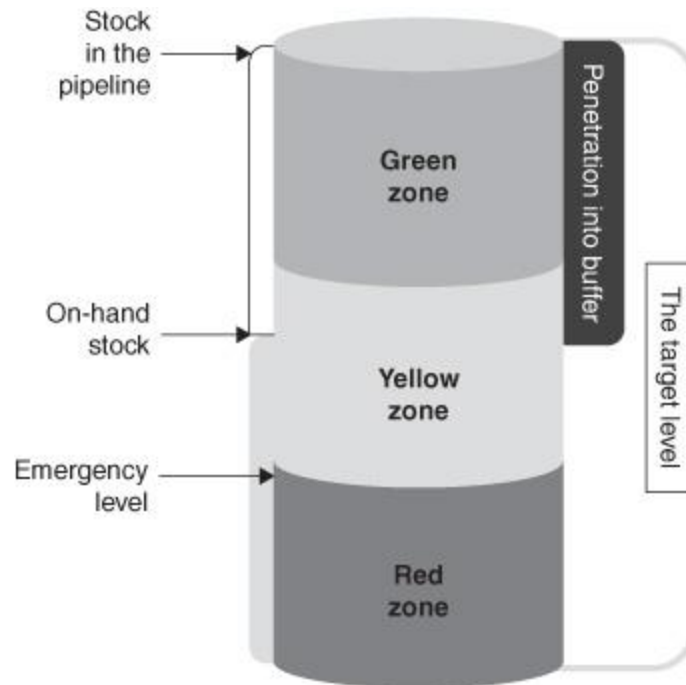


FIGURE 10-1 The structure of the stock buffer.

At any given point in time, the stock buffer is divided into the part that exists as finished goods on-hand and available for immediate sale, and the stock that complements the previous part to the full target level. Assuming we keep the target level intact, then the latter part is in the form of all the product components required for the finished goods to be equal to the target level. The part of the buffer that is not in the finished goods is called “penetration into the buffer” because that stock has not yet completed manufacturing and therefore is not currently available for immediate shipment. The buffer status is defined as the percentage of the penetration into the buffer. [Table 10-1](#) shows Order 1 with only 20 percent of the target inventory ahead of it in a status where 80 percent of the target is still on the shop floor. Therefore, its buffer penetration is 80 percent, which is greater than the 67 percent limit putting it in the red zone. When the penetration into the buffer is less than 33 percent, we are in green—actually too much finished-goods stock at

the moment. When the buffer is in yellow—buffer status between 33 and 66 percent—then the buffer state is truly satisfactory. (So we have at least one-third of the target in finished-goods inventory and the rest lined up in the shop for fabrication when needed.) Likewise, when the buffer penetration is equal to or above 67 percent, the buffer is red. The immediate message is expedite the order as you are about to stock out.

The priority rules are now clear: red orders should be expedited and should trigger management attention. Red orders definitely should have priority over all other orders, while yellow orders have priority over green orders.

Within the same color code, the decision of which order to do next is in the hands of the operators on the floor. The author believes that the buffer status, on top of the color code itself, is valuable information for the operator. If two red orders show up, one with buffer status of 70 percent and the other with 96 percent, it seems clear that one needs a very persuasive argument not to process the 96 percent order first. However, if one order is 70 percent and the other is 74 percent, then the real choice probably lies in other factors.

Generating Production Orders and the State of Capacity

The ideal situation is to generate new production for all items that were consumed the day before every day. What is the obvious negative branch of doing exactly that?

What could easily happen is that too much time is devoted to setups. Should we be concerned about doing too many setups? When at least one resource is losing too much of its protective capacity, then we definitely should care. The problem in losing the protective capacity is that the replenishment time grows longer and longer. Then, with longer replenishment times, more and more end products

would become red. When the number of red orders exceeds 20 percent, then the whole scheme of maintaining priorities loses its effectiveness and a significant number of shortages will occur.

The lesson we should keep in our mind is that *MTA requires a certain level of protective capacity*. We deal more with that issue because losing protective capacity might be caused by too much total demand (not just the demand for one item, but also the demand for the whole product mix). Right now, we do not wish to let too many setups be the cause for losing protective capacity.⁸ There are two ways to deal with the issue:

1. Dictating a minimum production batch. *The minimum batch is not part of the target level!* It comes on top of it. That means that once the inventory in the pipeline plus on-hand is less than the target level, a production order is generated, but its size equals a quantity at least equal to the minimum batch quantity. We may discover that the total inventory is above the target level, but it should be less than the target level plus the minimum batch.
2. Managing the capacity of the capacity constrained resource (CCR) and releasing new production orders only when it seems reasonable, the CCR would work on them soon. The concept of the “planned load” was defined in [Chapter 9](#); here we need to define the *planned load* for the specific environment of MTA.

Definition⁹: *The regular planned load for MTA is the summation of the derived load on the CCR of all the production orders already released that have not yet been processed by the CCR.*

Releasing orders only up to a certain limit of the regular planned load causes the release of new orders to be under

control as this procedure releases new production orders only up to that level where the regular planned load approaches the agreed limit. Production orders that were not released today will be considered again the next day. The regular planned load for the next day should be smaller by the amount of work the CCR has processed during the previous day and this allows more orders to be released.

What should the criteria be for choosing which production orders to release? The relative priority of each new production order competing for being released today is based on how much replenishment is required to reach a full target level of inventory in the shop floor (including the warehouse) relative to the target level. In other words, the production orders in the queue awaiting release are getting a buffer status, similar to the orders that were released. That status serves as the relative priority of the order. The orders with the highest status would be released first. Every order that is released updates the regular planned load. Once the planned load limit has been exceeded, the daily release of orders is stopped. The rest of the queue awaiting release has to wait until the next day, and as consumption continues, their relative priority will go up accordingly.

Let's demonstrate the mechanics of this procedure with the following example: Suppose the replenishment time is 5 days with 16 hours of CCR time every day ($5 \text{ days} \times 16 \text{ hrs/day} = 80 \text{ hours}$). A natural limit for the regular planned load is 80 percent of the replenishment time—64 hours. We use only 80 percent of the replenishment time assuming that this way the overall replenishment time, including the operations downstream of the CCR, would be easily maintained (the part downstream of the CCR usually takes much less time than the time to reach the CCR waiting for its turn and then being processed by the CCR). This procedure of release would avoid having too much WIP on the shop floor. Suppose that on a given day the planned load reaches 50 hours. Suppose 11 items are in the

production pipeline. The relevant information is shown in [Table 10-2](#).

As the planned load is already 50 hours and the limit is 64 hours, we can release up to 14 hours of work. Right now we need to release 19.3 hours—a little too much. The most straightforward method is to add from the highest priority, P10, then P3, then P1, P7, P5, P8. This would bring us to a total of 13.2 hours. The next, P9, would penetrate the 14 hours limit. Should P9 be released? This decision should be subject to the judgment of the person in charge (usually the master scheduler) and is not too critical. The main point is that P6, P2, and P4 would wait at least one additional day.

In most cases, a certain minimum batch is required. That must be considered in addition to the CCR load. *When a minimum batch is used, then the priority is determined just by the quantity required to replenish to the target level.* However, the load on the CCR needs to consider the size of the batch. For instance, if the target level is 100 and currently there are 49 on-hand and 50 in the pipeline (it could be that the 50 units are included in two production orders, not yet finished, somewhere in the shop floor, etc.). The replenishment to the target level requires only one unit, but the minimum batch is 25. The priority for releasing the next replenishment order is based on $1 \times 100/100 = 1\%$, but the time on the CCR needs to consider a batch of 25. When the load on the CCR and the relative priorities of the other orders will permit the release of this order consisting of 25 units to the floor, the time it'll take on the CCR is planned based on the load of processing 25 units. It may thus prevent other orders from being released on that day.

Product	Quantity to Replenish	Target Level	Priority (%)	Total Time on CCR ^a (hours)
P1	13	120	10.83	1.5
P2	3	95	3.16	0.8
P3	120	1000	12	2
P4	45	3000	1.5	0.8
P5	24	400	6	3.2
P6	114	3500	3.26	2.5
P7	100	1000	10	1.5
P8	100	2000	5	2
P9	33	750	4.4	2
P10	50	400	12.5	3
			Total:	19.3 hours

^aIncludes setup time.

TABLE 10-2 List of the Orders in the Queue Awaiting Release

Peak and Off-Peak Behaviors

At off-peak periods, there should be no blockages to the release of small replenishment orders to the floor. When the demand starts to pick up, then monitoring the replenishment order release starts to be critical because continuing to release small daily batches increases the number of setups, creating more blockages of the flow in the shop. The impact of this algorithm is to set a limit on the actual wait time to the CCR, thus limiting WIP to a degree that will allow orders in the shop to move at a speed that is in line with the assumed replenishment time. However, at the same time there are orders that still wait to go into the shop. This means the actual replenishment time is longer than the formal replenishment time. If that situation extends for a long time, there is real risk of losing control of the

availability commitment. The suggested behavior of prioritizing the release leads to a dynamic batching mechanism, where at off-peak the batches are naturally small (equal to daily sales) while at peak periods larger batches are used because of delaying the release of certain replenishment requests. This provides a little relief to capacity. In the short-term, this institutes the right priorities for the system in a way that saves setups on the CCR.¹⁰ We still need to examine the longer-term impact of monitoring the level of capacity and taking the right measures on time because enlarging the batches has only limited impact on the capacity and it could be that a true need to increase capacity is emerging.

Monitoring the Target Level Size—Dynamic Buffer Management

While the immediate process of fast replenishment of sales and following the right priorities in the shop floor have been covered, the next step is getting the right feedback to the planning stage. The most obvious planning decision is the determination of the target levels. The first initial estimation might not be adequate or changes in either the demand or the supply may have made a certain target level no longer adequate. What should the signals be that a specific target level is too high or too low? It certainly must be shown in the behavior of the on-hand stock. The algorithms for recommending changing the target levels are based on certain behavior patterns of the finished-goods stock and are called *Dynamic Buffer Management* (DBM).

Too Much Green—the Target Is Too High

When we hold target levels of several items too high, there are obvious negative effects. The direct financial implications and the risk of losing on investment are probably not too prominent assuming the initial determination of the level is not vastly wrong. *However, holding too high inventory means we replenish when there is no real need.* Therefore, it has a direct capacity impact that at off-peak time might not be problematic, but at peak times could be critical.

The most obvious signal that a buffer is too large is that it is too often and too long “in the green zone.” Stock buffers are not supposed to be in the green for too long. It means that the relationship between supply and demand does not call for such a large buffer. We call the situation “*too much green*”—*a signal that the buffer target is too high.*

Let’s define a parameter called the “green check period” that whenever an item spends time continuously in the green it is recommended that the target level be decreased. The recommended default time for the green check period is twice the replenishment time. The point is to be reasonably conservative. It is not desired to reduce the buffer (the target level) and after a short time to increase it again. Frequently, holding a little too much inventory is preferred to holding too little.

Once the target level is reduced, it is natural for the current on-hand stock to be *above the new target level.* No checking and definitely no decision on further reduction of the target should be considered when the current on-hand stock is above the top of the green, which is equal to the target level.

Once it is decided to reduce the buffer, the next obvious question is by how much? Goldratt suggests (Strategy and Tactic tree MTS to MTA 2008, entry 5.112.1) to reduce the target level by 33 percent. The topic of by how much to increase or decrease buffers and when to refrain from doing

so is a worthy topic for discussion. We'll deal with it later in this chapter.

Too Much Red—the Target Is Too Low

“Too much green” is the signal that the buffer target is too high, and *too much red points to a buffer target that is too low*. That said, we would like to be a little more precise about increasing the buffer. Spending a lot of time near the top of the red looks bad, but may not be bad enough to propose increasing the buffer. In addition, it could be that every time the buffer turns red it is replenished quickly, but soon turns red again. This might be a signal that the buffer is not high enough to prevent the risk of shortages.

The idea is that both the amount of time an item spends in the red and the depth into the red are relevant signals to increase the buffer. The algorithm that emerges is that every time there is a penetration into the red, the depth of the penetration, expressed by the number of item units below the red level, is recorded. If within the time frame of the replenishment time the summation of all the recorded penetrations is equal to or greater than the size of the red level, then a recommendation to increase the buffer is given. In other words, if during the span of the replenishment period the penetration of red equals the entire size of the red buffer, it is time to increase the buffer size.

Once the target level has been increased, the specific item will definitely be in the red. The increase in the target causes a new replenishment production order to be released. It certainly will take time before the new buffer size will be stabilized. Before that, there is no sense in deciding to increase the buffer again. The point here is to refrain from hasty decisions until the impact of the previous increase has been noted.¹¹ Thus, the algorithm calls for a

“cooling period” where no re-evaluation of the penetration into the red would be done. The natural time for the cooling period is one replenishment time. Therefore, it takes one replenishment time to possibly discover the buffer should be increased and another period of replenishment time until such a check should start again.

Discussion: Issues with DBM and by How Much to Increase/Decrease the Targets

The first topic in this discussion has been by how much to increase/decrease the buffer. From this question, some additional questions might be raised, such as what are the immediate ramifications for such a change, and due to them when should such changes be avoided? Shouldn't the increase of the buffer be subject to a forecast, which predicts how much the demand would grow?

In practice, the sales of one item at a specific location are too chaotic to truly support a good prediction of the quantity. However, the trend of the sales can be predicted, so we should know whether we need to increase or decrease the buffer, and decide rather arbitrarily about the size of the change.

We discuss here the behavior of sales from the manufacturer's viewpoint. In other words, wild fluctuations are less common at the manufacturer's level than at a specific store. The question is whether we'd have a better answer for the manufacturer than the arbitrary guideline that says whenever a clear signal is noted that the buffer is not adequate, change the buffer by 33 percent or any other fixed ratio that seems appropriate.

Note that the BM signal is impacted by the combination of demand and supply. When the demand goes up, the idle capacity decreases and the replenishment time gets longer.

Do we know how that is going to affect the right size of the buffer?

This author's inclination is to accept the premise of having an arbitrary number for buffer increase or decrease. However, for the shop floor a decision to increase the buffer by 33 percent looks to this author like it creates too many waves in the general flow. A buffer increase of 20 percent and a buffer decrease of 15 percent look more appropriate for the shop floor. The demand from a manufacturer usually has much less fluctuation than the sales of a store, and thus the changes in the buffers could be smaller and still be able to match the trends.

Another question is what are the appropriate conditions for increasing a buffer? When the buffer is increased, the whole amount of the increase is released to the shop floor as one production order. This relatively large production order comes on top of the regular replenishments that are following the actual demand. If the current load on the CCR is high, then the last action we should take is to release another large order to the floor.

Actually, when there is high load on the CCR, it could easily generate a recommendation to increase the buffers of several items. If production management takes the recommendation to increase the buffers, then a substantial additional load will be added to an already high-load pressure. This extra load might cause more items to penetrate the red zone for too long, causing even more items to penetrate the red for too long. It can easily turn into a vicious cycle! Thus, the point is to allow buffer increases only when there is no immediate load pressure, or when additional capacity can be used for it.

Of course, buffer decreases are easier to make and they reduce the load pressure. However, if the recommendation to decrease the buffer is not justified, then some time later we would see a recommendation to increase the buffer;

then, depending on the total capacity situation it might be difficult to do.

As we have seen, certain supply problems can cause “false” recommendations to increase the buffers. By “false” we mean either that the problem observed by buffer management is just a rare statistical fluctuation or that while there is a real problem with the current specific target levels, the targets should *not* be increased at that time. This is the case when a temporary lack of a specific raw material is causing the end item to go into the red without being able to replenish it very soon. In such a case, an increase of the buffer would not help anything. Once the missing material shows up, the dilemma of whether to increase the buffer or not can be dealt with. The main point behind the dilemma is whether the manufacturer may want to protect itself from a future lack of materials (because it happened this time) by building a relatively high level of stock. The more sensible alternative is to increase the buffer of the specific raw material and settle the issue that way.

DBM is vital for getting the right signals about the validity of the buffers. In manufacturing, the author highly recommends that *any actual decision about changing a buffer should be judged by the human mind and in no way be left for a software package to dictate*. Asking and understanding why a buffer is continually in the red or green should be done prior to increasing or decreasing the buffer target. In a distribution organization, the number of buffers makes such a human judgment on any buffer change very difficult.

The power of DBM comes from judging the combination of demand and supply. However, the author thinks that for the sensitive decision of increasing or decreasing a finished-goods stock buffer in manufacturing environments, a focused analysis of both the demand and the flow in the production shop floor, pointing out to possible critical changes in their respective behavior, and checking possible

shortages of materials should be key in the decision. Such an analysis should be done fast, based on focused information that should be part of the information system, so that the decisions can be made quickly. Right now, such an analysis is not a part of the known TOC solution for MTA.

The Role of Protective Capacity and the Usefulness of Maintaining a Capacity Buffer

The need for protective capacity to maintain availability has already been mentioned in this chapter. A special problem of MTA is that the commitment to the market cannot be conditioned by a total amount. You certainly can tell your customers that the commitment is to maintain availability up to a certain level of one-time demand. This way you protect yourself from excessive one-time demands. For an item whose target level is 100 units, a demand of 30 units coming from one customer at one time is already problematic, and a one-time demand for 60 units cannot always be answered even when the MTA procedures are done according to the book. Therefore, it would be wise to tell customers that the commitment for availability for that item is limited to up to 15 units per customer at one time. The idea is that one-time demand of up to one-half of a zone, one-sixth of the whole target level, is still acceptable. Customers that wish to draw at one time a larger quantity should issue an order to be supplied at a certain quoted lead time, like any regular MTO type of orders.

Nevertheless, such a limit on the commitment is not capable of addressing a 20 percent rise in the total demand for all the products at the same time. After all, it is not the responsibility of any single customer to look at the total demand of all customers.

When we dealt with MTO, we established a way to deal with too much demand by being able to quote longer lead

times than the standard lead time. This method of quotation smoothes the load and allows good utilization of the CCR.

In MTA, we do not have a way to restrain the demand according to capacity. Does it mean we should have to maintain enough capacity at all times? Well, it is certainly possible to sustain a peak of load for a limited time because having enough inventory on hand smoothes the impact of a temporary lack of capacity. The effectiveness of BM priorities dedicates the limited capacity to those products that need it most. However, such a peak period cannot continue for too long without affecting availability. Thus, the biggest risk to the good performance of the TOC MTA methodology is growing market demand, which requires more capacity than the CCR is capable of handling.

The planned load is a worthy tool to judge the required capacity based on the demand at hand. However, in order to judge the capacity at hand we must include in the planned load all replenishment orders. As you may recall, the regular definition of planned load for MTA takes into account only the replenishment orders that are released into the floor. Ideally, all the replenishment requests have been released to the floor, but when capacity is temporarily limited, some replenishment orders are delayed because they have lower priority and the CCR would be busy anyway processing orders that are more pressing.

Therefore, in order to monitor the overall capacity status one must run a somewhat different planned load; let's call it the *full planned load that includes all production orders and also replenishment orders that are not released yet*.

What we get then is the time it takes for a new production order, just released, until it is processed by the CCR. Ideally, we need it to be not more than 80 percent of the formal replenishment time. The other 20 percent represents the time required after the CCR to complete the order. When the planned load is longer than 80 percent of the replenishment time, it means the actual replenishment time is longer and

thus there might be a threat on the availability if that situation continues. Therefore, if this is just a peak for a short period of time, then the system has a good chance to stay stable. However, if the market demand continues to grow, then the threat would become a reality.

If there is a reasonable assessment that the demand is truly going up, then the conclusion should be to increase the capacity as soon as possible.¹² Of course, we mean to increase the capacity on the CCR, but any elevation of the CCR should initiate an analysis of whether another resource would become a true CCR, so we might consider increasing the capacity of that resource as well.

Increasing capacity is an investment, so we had better be confident that the sales are truly growing. In addition, we need to know which resources should be elevated. We know the CCR requires more capacity, but many times we have less good information on the other resources. As we'll see later, certain feedback from the floor might help us to pinpoint the resources that will need extra capacity once the demand goes up. However, additional study of other resources that might require more capacity may be needed.

There might be another way. Suppose there is a certain amount of capacity that can be purchased quickly at will; for instance, calling for overtime or even adding an extra shift. There are cases where the shop floor already works 24 hours a day, 7 days a week. Even at that rate, sometimes night shifts are not fully manned. Another way of purchasing capacity at will is by outsourcing. What is typical of all these cases is that the extra capacity costs additional money every time it is used. Moreover, to preserve that amount of capacity one needs to use it from time to time, otherwise it won't be available when the need arises. Suppose that for a whole year no extra shift was called; how easy will it be to organize it? It is not certain there is enough labor to staff it

and if it is possible on paper to gather the required people, do they really wish to work an extra shift?

This author defines *capacity buffer* as a quick means to purchase additional capacity that is truly available on reasonable notice. It is *a buffer to protect the ability of the company to commit availability and truly meet the commitment*. As a true buffer, the level of use of the buffer signals the level of pressure the system is under. The use of the capacity buffer should be initiated by two different parameters: the full planned load (when it is larger than it should be) and the number of red orders relative to the average number of production orders. The full planned load approximates the real load and it depends on the accuracy of the data. However, when the full planned load is growing beyond the previous limits, one must deduce that higher than usual red orders will follow. Thus, additional capacity can be planned based on the early warning of the planned load, or wait for the emergence of red orders and then add the required capacity. Following the increase in red orders has the advantage of knowing which resources are required for expediting. This is very valuable information for a decision on investments in capacity.

The capacity buffer behaves like every buffer. The use of it could easily conform to the green-yellow-red mark. However, the capacity buffer should be in the green most of the time—meaning less than its potential is actually used. When the regular use is already in the yellow, then its function as a buffer is already compromised to a certain degree.

The Process of Ongoing Improvement (POOGI)

The fourth concept of flow¹³ (Goldratt, 2008) speaks about establishing a focused process of balancing flow. It certainly fits well with the way BM functions—giving the right

priorities to what should be done now. However, balancing the flow should be taken seriously for the longer time frame as well. In other words, we must have a focused mechanism to identify specific areas where an improvement would really improve the overall flow.

Again, BM supplies certain basic information. In MTO every time an order penetrates into the red zone the user should enter a “reason” from a table, so monthly analysis can be done to pinpoint the most frequent reason and see what can be done to eliminate it. On top of the list of reasons, it is possible to collect the whereabouts of the production order when it turned red. The assumption is that in most cases for a resource that causes long delays there will be many times when production orders will turn red while waiting for that resource.

In MTA, entering the red has three possible causes. One is too long of a delay in order release (lack of materials or too high-load pressure). Two is too slow of a flow in the shop floor that caused the on-hand stock to penetrate the red level. Three is high sales within the last day or two. When we look to balance the flow, the third cause is not relevant. Only causes for relatively long delays are relevant for such a process.

The new suggestion by Dr. Goldratt (Strategy and Tactic tree MTS to MTA, 2008, entity 5.113.2) is to register any delay that is “too long” in a work center. The suggested definition for “too long delay” is one-tenth of the formal replenishment time. Registering such a delay does not ensure entering into the final Pareto list for picking the highest one area and trying to improve it. The other condition is that the delayed order would eventually enter into the red. Only then does that occurrence enter the Pareto list.

The feedback process requires reporting when an order arrives at any work center in the list of “to be watched” and reporting when that order has been completed at that work

center. In order to implement this process, the organization needs both the appropriate software for reporting and the discipline of the operators to report it accurately.

Generic Issues in MTA

MTA for Components

The main essence of MTA is the combination of marketing and operations. Let's now consider the possible value of producing certain common components to availability. Here we are speaking of common components used in MTA finished goods for customers. Like in the original meaning of MTA, we better have a strong commitment to maintain excellent availability of those common components. The point is that in planning the production of the end products, either to order or to availability, it would be possible to rely on the availability of the common components. The value, when it is truly applicable, is a major cut in the production lead time.

The main reason for maintaining stock for common components is to shorten significantly the response time to the market. Another reason has to do with minimizing the significant setup times that are sometimes involved in producing basic parts/materials used in many end items. For instance, in many V-plants, such as plastics or paper, there are a few base materials, which are used in many end items. The primary operation in preparing the base materials (like an extruder) often has very large setup times. The operational problem stemming from the long setups is that many urgent orders for a certain base material emerge all the time, as any demand for one of the huge number of end products creates a demand for one of

the base material products. The longer is the wait time for the primary operations, the more end-items and the more production orders enter the red zone, thus putting pressure on the primary operation. Urgent requests (usually for a small quantity of one base material) make it difficult to keep the minimum batch size required to keep that primary resource from becoming a bottleneck. The straightforward solution is to produce those base materials to availability, and then the only urgent replenishment is when the stock of one of the base materials is penetrating the red.

Producing base materials or components to availability splits the entire manufacturing floor into two different environments separated by stock buffers. Both are planned in separate runs, even though one environment feeds the other.

The planning and the BM for MTA of components is the same as for MTA of end items. Note two important points:

1. In order to provide a smooth transition of an item from order to availability, the initial stock must be in place.
2. Every item should be defined as either to order or to availability. If there is a need to produce the same item both to order and to availability, then define the two separate Stock Keeping Units (SKUs), one for orders and the other for availability. This point will be discussed later in the section on mixed environments.

Which Items Fit MTA and Which Fit MTO?

It is quite clear that not every item should be managed to availability. One factor to consider is the level of the demand fluctuations. Here is a simple graphical representation of a typical MTA item demand versus a

typical MTO item demand. Let us start with a typical MTA shape in [Fig. 10-2](#).

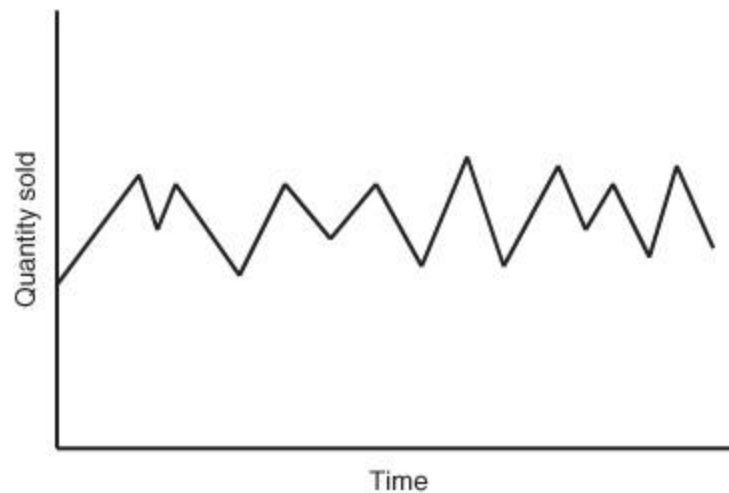


FIGURE 10-2 Semi-continuous behavior of sales for a typical MTA item.

Sales for an item that nicely fits managing to availability is having a spread of daily consumption that is less than the average daily sales (Coefficient of Variation less than 1). That also means that on most days there are some sales. Such a spread allows for frequent and fast replenishment and the on-hand stock would stay mainly in the yellow.

Other items might have a very sporadic demand that fits a MTO pattern. Most days there is no single sale, but there are days where the clients purchase relatively large amounts. It could look like [Fig. 10-3](#).

In order to manage such an item in MTA, there is a need to hold a very large target level and in a substantial amount of the time the on-hand would be in green, which also means that DBM should not handle very well such an item. While the clients might wish to include such items for immediate availability any time they need it, the characteristics of the item are such that everybody else would have difficulty to make it available at all times. Thus, if a relatively short

response is offered to the clients they would, most of the time, accept it.

The same is also true for managing common components to availability; the criterion depends on the shape of the consumption curve and the size of the spread.

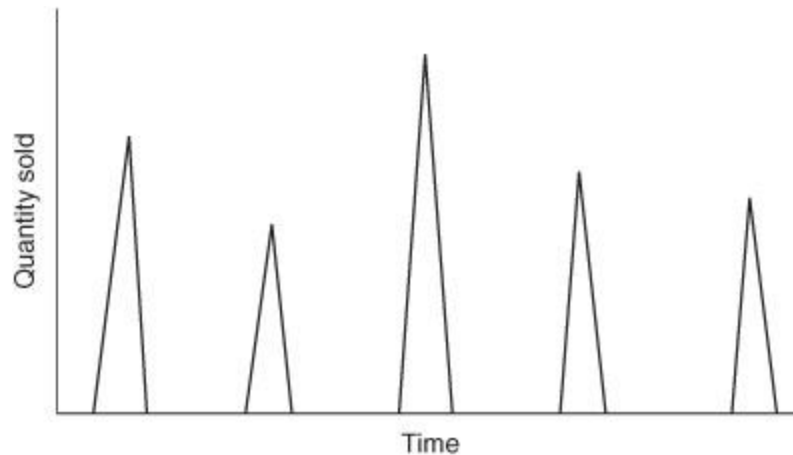


FIGURE 10-3 Sporadic demand that should be better managed as MTO.

[Chapter 11](#) is dedicated to the way TOC handles distribution, Amir Schragenheim offers a way to consider the *return on investment* (ROI) in carrying stock of an item. When applied to manufacturing, the same parameters still apply, and the spread of consumption is reflected in the target level necessary to maintain availability. This is a somewhat more detailed approach for consideration.

Vendor-Managed Inventory (VMI)

The whole point of MTA is to offer new business opportunities due to the extra value given to the clients, which competitors would find hard to imitate. A more particular opportunity is to ask a relatively large business client to take responsibility for the level of stock of the manufacturer's items, at the client site. This kind of

business relationship is known as vendor-management inventory (VMI).¹⁴

VMI is not the invention of TOC. It is well known because some ultra-large organizations force it on their small- to medium-size vendors. Thus, it is justifiably a win-lose kind of relationship. The vendor has to comply with whatever the big guy tells him to.

Understanding how to run MTA effectively raises the business opportunity where a vendor could offer to a client a desirable alternative that is typically forced on clients. We won't detail here the business relationship and how such a deal could be win-win for both the vendor and the client. From the logistical part, we need to make the distinction between standard items that are sold to many clients and items that are dedicated to the particular client for whom the offer is given.

Items that are sold to many clients should be managed according to MTA up to the plant warehouse, and thereafter by the distribution solution.

When the items are dedicated to one client for whom VMI is already in place, then there is no point in maintaining stock of these items at the plant warehouse. It could be the case that between the production and the actual shipment there is a practical need to store the items for a day or two. However, the real storage and the focus of BM should be on the client's site.

The replenishment time for VMI should include the transportation time. The characteristics of the transportation could be important because it is difficult to expedite a shipment once it has been sent. VMI is much more effective when the transportation time is much shorter than the manufacturing time. This seems to be true in the majority of cases.

Mixed (MTA and MTO) Environments

As it was already mentioned, not all items should be managed to availability, with the alternative of offering short delivery time to selected items (as MTO). It is clear that in many cases a mixed environment is advised.

There could be two different meanings to the term *mixed environments of MTA and MTO*. The first is that there are items that are strictly produced to availability and others that are strictly produced to order. The other possible meaning is that several items have demand both for immediate delivery and (usually for large quantities) for future delivery at specific dates.

The generic problem is how to manage an environment with two different types of buffers: time and stock. There are several sophisticated ways to do it, but we want the simple and effective way. Having the same items under two different management systems is too complicated to be truly optimal in reality. Our clear suggestion is to separate the item SKU identifier of the MTO from the item managed to MTA, even though for all practical purposes they are the same item. When the MTO orders for an item, usually managed as MTA, are treated as orders for a separate item, then there might be a case where the MTO order is expedited even though there is enough stock on-hand to deliver it. A similar case can happen when it seems we need to expedite the production orders for the MTA item, but an MTO order has been finished long before its due date and thus there is enough stock to cover the immediate demand. Our advice is to ignore those cases and simply expedite even though there may be another solution for it, or the production manager can take it on as his own responsibility to stop expediting and make the necessary shift between the two identities of the item (having two different SKUs for the same physical item).

This leaves us with managing MTA items to the side of MTO items. Before we state the solution, we recommend

addressing the flawed perception of priority in reality. Put yourself in the shoes of an operator who has to choose between two orders: one is an MTO order, with a specific client willing to pay, the other is for stock, meaning we do not know when a client would buy it directly from stock. It seems that the MTO order has a clear priority because it means “Throughput now,” versus “possible Throughput sometime in the future.”

Suppose the MTO order is to be delivered in three weeks, while the MTA order is for a product that is currently short at the finished-goods warehouse. Would you now also prioritize the MTO order? If not, then what is the rule?

Another perspective is proposed. In MTA, the company is offering a commitment to provide availability. The same commitment is given to a client for the MTO order by specifying a date where the order would be completed. Thus, the issue of priorities means: *what are the priorities to follow for the best chance to meet ALL of our commitments?* BM yields the green-yellow-red priorities and the claim is that even though both have different sorts of buffers, the meaning of the green-yellow-red, and even the meaning of the buffer status, is exactly the same. When an operator is facing orders with various buffer statuses, he does not need to know which one is MTO and which one is MTA. The color code is the main priority mechanism, with the buffer status as additional, more detailed information.

There is one problem concerning the mixture of MTA and MTO: managing capacity. MTA is stricter in its requirements of maintaining protective capacity of the CCR because in MTO we have the flexibility of quoting longer lead times when necessary. When the amount of MTA relative to the MTO is relatively small, say approximately 15 percent MTA and 85 percent MTO, then reserving 15 percent of the capacity for MTA and basing the time quotation of MTO on 85 percent of the available capacity is an acceptable solution. For all other cases, we suggest taking the MTA

capacity management as the overall rule. It means MTO would be handled based on time quotation of standard lead time, assuming it is always possible because there is enough capacity at hand. Having the capacity buffer in place (the ability to increase capacity easily and rapidly) is an excellent way to draw the maximum capacity from the internal resources and using the capacity buffer whenever needed. More on this approach can be found in Schragenheim, Dettmer, and Patterson (2009, [Chapter 7](#)).

Dealing with Seasonality

Seasonality poses obvious problems in managing stock in general. [Chapter 11](#) in this Handbook deals at length with this issue, including the dilemma to choose between forecasting the demand and simply following DBM.¹⁵

In this chapter, we highlight the acute problem for the manufacturer. Managing capacity is a special problematic area in manufacturing, often for the wrong reasons, such as achieving high efficiency of every single resource. From the pragmatic viewpoint in TOC, managing capacity is still an issue—making sure there is enough capacity to meet the demand.

In most cases, when the term “seasonality” is used, the meaning is peak demand within a certain period of time. Such a peak of demand could take several months or just one or two days. There is a clear difference between these extremes; a very short peak means no replenishment could take place during the peak. This case will be dealt with as an MTS later in this chapter.

The capacity problem with seasonality¹⁶ is that *within the peak the total demand might require more capacity than the CCR has*. Such a situation would definitely reduce the on-hand stocks and all one could do is try to prioritize. For a

short while, it could be good enough, but for a longer period of time, it would be disastrous.

Increasing the target levels before the start of the season is only a partial solution to the capacity problem. If there would be a real lack of capacity during the peak, and if the peak is not very short, then shortages will certainly occur.

Solving the capacity issue requires investment in capacity and materials before the start of the peak. The direction of the solution is to create enough inventories of several fast runners to cover the demand for those items throughout most of the whole peak. A valid way to do it is to *forecast the minimum quantity to be sold through the whole peak of several fast runners* and produce this amount prior to the peak. Not having to replenish those fast runners every time there is a sale would save precious capacity that will be used to replenish the other items. The reason we suggest doing this only for several of the top fast runners lies in the characteristics of fast runners to have less spread of the future sales. Even if some of the inventory is left after the peak, it will still be sold.

Problematic Environments for MTA

The replenishment solution for manufacturing is dependent on the priority mechanism in the execution phase. Not just on the priority itself, but even more on the capability to expedite. Environments with longer replenishment time and where expediting is either impossible or very difficult have to compensate for the lack of flexibility by maintaining more stock and by frequent replenishments. Even when the latter is possible, there is a real problem in achieving good availability.

Consider the case of sequence-dependent setups or even just very long setups coupled with a long list of items to be produced. For example, a production line of paint produces

12 different colors. Each of them has three to five different variations of the paint. In producing paint, the length of the setup (mainly cleaning the line from any residuals of the former paint) depends not only on the next color, but also on the previous one. When you keep the sequence that goes from light colors to the dark colors, the total setup time is much less than trying to produce according to the sequence of the real needs of the market. Hence, the production line has to stick to the preferred sequence and thus produce the whole cycle (certain slow items might be skipped from time to time). Suppose that the whole cycle takes 21 days (3 weeks). This means that the replenishment time is 21 days.¹⁷ What should the target level be? The “maximum” consumption is within 21 days, so the average sales within 32 days seems appropriate enough.

However, note that it means every item is replenished only once in 21 days. So, if the sales are higher than expected, toward the 20th day the on-hand stock might be penetrating deep into the red and there is nothing we can do about it. Deviating from the sequence might have too much impact on the capacity, maybe even turning the production line into a bottleneck.

The only remedy to this state is to hold much more stock. If the target level is twice the average of sales within the cycle time, then most of the time the on-hand stock will be in the green and toward the end of the cycle, it'll be in the yellow. This means the “too much green” parameter needs to be long enough to prevent a false recommendation to reduce the target.

Dealing with the more problematic environments should highlight how good the TOC solution is for most other environments.

MTS That Is Not MTA

There are certain cases where it makes perfect sense to use MTS, but this is not coupled with a commitment to maintain perfect availability. We can identify two categories of such cases:

1. The reasons for MTS come from capacity management and not for ensuring availability.
2. The organization is trying to provide a certain level of availability but cannot, or even does not wish to guarantee it.

Let's analyze the aspects of the two categories. The first one has already been demonstrated by the seasonality approach where sometimes high stocks, well above the target levels, are used on a few fast runners to free capacity during the season itself. Preparing for any peak of demand that requires capacity above capabilities forces the capacity planners to look for MTS even in MTO situations. Of course, if the MTO orders are all fully customized to the requirements of the clients, this is not possible but it could be that stocking some components would still relieve the pressure of capacity on the CCR.

The other category is typical of situations where the possible demand fluctuations are too high to provide excellent availability or the surplus inventory is very expensive to hold. In such situations, the marketing approach could be that we do not promise availability, so if you really want to buy, be fast! Examples of where availability cannot be guaranteed are:

- Launching new products, especially new innovative products.
- Promotions where a peak demand is anticipated, but perfect availability should not necessarily be offered to

the clients.

- A short demand peak where replenishment within the peak is impossible.
- Products with very short shelf life where the truly varying costs of producing them are high relative to the Throughput.

There is no point in committing to availability in such cases. Still, providing availability can be the main force behind the decision about how much to produce, but full commitment seems risky.

What should the process of managing MTS look like? Two distinct problems for MTS are not present in MTA. First, how do you decide how much to make to stock? The decision should be based on a forecast that recognizes both the reasonable minimum and maximum sales within the reliable supply time and being aware of the damage of shortages and the damage of surpluses.

Second, how do you prioritize a production order for stock and not for availability? Checking the state of the downstream inventory relative to the target level does not make sense. It seems to us that when the stock is intended for an anticipated peak of sales, setting a date for completion and treating the order as an MTO makes more sense. The date in such a case is not really artificial because the stock is targeting a certain peak demand event.

In the cases within the second category but where no special event is the trigger, the marketing approach of creating an atmosphere of being “hard-to-get” makes the replenishment technique still valid. However, the target levels are intentionally low and the expectations are to be in the red (or even in the black) most of the time. The expediting efforts might be much more restrained and most of the DBM recommendations are not going to be granted.

Implementation Issues

Some of the implementation issues have already been addressed within this chapter, especially determining the initial values of the target levels. Certainly, the process of deciding upon the initial values of the target levels should be as short as possible. There is no need to be precise—a very rough estimation is more than enough.

An important point is the initial calculation of the targets. One should assume the replenishment times to be much shorter than the current one. In many cases, cutting the current replenishment time by half is a good initial guess for setting the target levels.

Buy-in issues won't be discussed in this chapter.¹⁸ The two problematic areas discussed next are included due to the potential problems that might occur and not be fully understood.

Moving from MTS or MTO to MTA

The move toward MTA can be done from either MTO or MTS or a combination of the two. The move from MTS to MTA might be viewed as relatively easy because there should be stock in the system, both finished and in production, that is not already dedicated to customer orders. Most of the time the amount of available stock found in the system is larger than the target level. The problem comes when there is no stock that is not already dedicated to specific customer orders.

When the shop currently operates under strict MTO, then there is definitely no WIP or finished goods that are not already dedicated to customer orders. The practical action is that enough stock must be prepared. The preferred way to start the replenishment is with full buffers—the finished-

goods stock buffer level is at the target. Only when the actual stock level allows for replenishment should the switch between MTO to MTA be performed. If that change were done too early, chaos would occur in the shop floor.

The problem is that building stock has to be done on top of continuing to supply to order according to the current rules and sales volumes. This means producing to stock (later it'd be for availability) on top of running the production order required for actual customer orders. The practical action is to use only excess capacity to build the required stocks. It would take time according to the available capacity. The best way to physically build finished-goods inventory buffer on top of the regular demand is to generate dummy MTO orders for the stock building using a large time buffer in finished goods so that it won't turn red too soon. The chosen date would give realistic expectations for when the availability marketing offer would be ready for launch.

Software Considerations

MTA places a lot of emphasis on the proper management of the shop floor. On top of the production side, at least two additional functions must be connected and performed well—marketing and sales (mentioned here as if they are the same function) and the purchasing of raw materials.

The marketing side needs to develop the offer to the clients, including setting the right expectations, so customers know about any limits to the quantity that can be purchased at one time. Sales needs to know when to try harder and when to find ways to restrict the demand based on the planned load.

The TOC buffers in both MTO and MTA assume perfect availability of raw materials. The generic rules of maintaining stock, explained in both this chapter and the

next chapter on distribution, are applicable for managing the stock buffers of the raw materials.

It is the natural role of software to connect all the pieces to show one holistic picture. However, that natural role is seldom what one really gets. Most ERP software packages do not really display one cross-functional picture. The current situation with TOC is that while the general direction of looking at the performance of the whole organization and pointing to the weak links does exist, it has not yet been converted fully into software specifications.

Another role of software is to focus the attention of the decision makers on what really matters. It is not done currently in the vast majority of the software packages.

Software has a critical role in communicating and instituting processes and terminology. This does not mean that the current ERP packages institute the right ones. TOC challenges many of them. However, the power of the software cannot be ignored. This highlights the need for TOC software to institute the right processes and to provide the relevant information to the various levels of the decision makers.

As no current ERP package is built upon the TOC principles, and given the basic difficulty of implementing/revising such software, the practical options are either to force the TOC procedures within the existing ERP or any other information system or to link add-on software to the ERP.

Concentrating just on the MTA requirements, there are five different areas of need for TOC functionality within the information system (IT) of the organization.

1. Generating production orders based solely on replenishing to a defined target level.
2. Generating the green-yellow-red generic priorities for every single production order. The buffer status should

be considered as a “bonus,” which is nice to have but not a must.

3. Using DBM to recommend changes to the target levels.
4. Monitoring capacity through planned load, including being able to recommend what replenishment orders to release.
5. Providing managerial reports, including POOGI, and monitoring the number of red orders and historical behavior of the planned load.

The first area can be quite easily forced on the MRP/ERP. Still, the terminology of target levels or buffers will not be included unless a more massive development is done. It means the people handling MRP need to understand the TOC logic and the terminology well to keep the ERP updated. The real difficulty lies in forcing the MRP/ERP to determine the green-yellow-red priorities. All ERP packages assume every work order should have a date, but the TOC logic is quite different.

Dynamic buffer management is another module that cannot be easily supported within the ERP itself.

One might be able to create a variety of reports of the planned load within the ERP, but it is far from being straightforward.

All of these points are relevant for the development of the organization’s approach to the software side. Every implementation should consider the options for software as an integral part of the implementation.

References

- Arnold, J. R., Chapman, S. N., and Clive, L. M. 2008 *Introduction to Materials Management*, 6th ed., Prentice Hall.
- Blackstone, J. H. 2008. *APICS Dictionary*. 12th ed. Alexandria, VA: APICS.
- Goldratt, E. M. 2008. Strategy and Tactic tree. Consumer Goods Make-to-Stock (MTS) to Make-to-Availability (MTA) S&T, Level 5, September 2008.
- Goldratt, E. M. 2009. "Standing on the Shoulders of Giants." *The Manufacturer*. June. http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants (accessed February 4, 2010).
- Goldratt, E. M. and Cox, J. 1984, 1993. *The Goal*. Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. and Fox, R. E. 1986. *The Race*. Croton-on-Hudson, NY: North River Press.
- Schragenheim, E., Dettmer, W. H., and Patterson, W. 2009. *Supply Chain Management at Warp Speed*., Boca Raton, FL: CRC Press.
- Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>.

Suggested Reading

www.inherentsimplicity.com/warp-speed is a site that allows downloading of the MICSS simulator including analysis files and more related materials.

See also Schragenheim, Dettmer, and Patterson. 2009. *Supply Chain Management at Warp Speed*. [Chapters 6](#) and [7](#) are especially relevant.

About the Author

In the last 25 years, Eli Schragenheim has taught, spoken at conferences, and consulted in more than 15 countries, including the United States, Canada, India, China, and Japan. He has also developed software simulation tools especially designed to experience the thinking of TOC, and consulted with several application software companies to develop the right TOC functionally in their own packages.

Mr. Schragenheim was a partner in the A.Y. Goldratt Institute and he is now a Director in The Goldratt Schools.

He is the author of *Management Dilemmas*. He collaborated with William H. Dettmer in writing *Manufacturing at Warp Speed*. He also collaborated with Carol A. Ptak on *ERP, Tools, Techniques, and Applications for Integrating the Supply Chain*, and with Dr. Goldratt and Carol A. Ptak on *Necessary But Not Sufficient*. In March 2009, a new book titled *Supply Chain Fulfillment at Warp Speed*, with William H. Dettmer and Wayne Patterson was published. The new book contains much of the new developments of TOC in operations.

Mr. Schragenheim holds an MBA from Tel Aviv University, Israel, and a BSc in Mathematics and Physics from Hebrew University in Jerusalem. In-between his formal studies, he was a TV director for almost 10 years. He is a citizen of Israel. The author's personal email is elyakim@netvision.net.il. Readers should feel free to write to Eli Schragenheim to discuss matters related to the chapters on MTO and MTA.

CHAPTER 11

Supply Chain Management¹

Amir Schragenheim

Introduction: The Current Practice of Managing Supply Chains

It is Wednesday afternoon. I am entering the grocery store and want to purchase some green peppers. However, they don't have any in stock. I can't find any good looking tomatoes either. I'm continuing to the Office Depot store. I heard great reviews about a new mouse that Microsoft issued and I would like to get one. However, I come to an empty shelf with only the item description stating "out of stock."

How many times have you gone to a shoe store to purchase a pair of wonderful shoes you wanted but they didn't have any in your size?

Why don't stores keep the right stocks to fulfill their demand? They seem to have a lot of stock. Why can't they do this simple task right?

Supply chains in our modern age operate in a way that seems to make a lot of sense. Manufacturers have robotic machinery to automate processes; many manufacturers have already installed new state-of-the-art Enterprise Resources Planning (ERP) systems to help them manage their shop floors.

Distributors and manufacturers have very sophisticated forecasting software to predict exactly how many items will

be sold of each product and even each stock keeping unit (SKU).² Therefore, they should know how many units they would like to send the retail stores (consumption points) and when.³

How is it that organizations still experience problems in managing the supply chains? Is technology not enough?

Problems with the Current System

Typical problems⁴ of supply chains are low inventory turns, high inventory investment, stockouts causing lost sales at some locations and at the same time excess inventories of the same items at other locations, high inventory obsolescence, lack of responsiveness to customer needs, etc. Let us examine some potential causes of these problems.

The Natural Tendency for Push Behavior

The vast majority of supply chains today are push systems. A *push system* in the *APICS Dictionary* (Blackstone, 2008, 112) is defined as “. . . 3) *In distribution, a system for replenishing field warehouse inventories where replenishment decision making is centralized, decisions are usually made at the manufacturing site or central supply facility.*” (© APICS 2008, used by permission, all rights reserved.). Given this definition, the centralized position in the supply chain is the manufacturer that supplies his regional warehouse or consumers directly or a distributor that purchases items from several manufacturers and distributes them to his regional warehouses or directly to the customer.

What is the *manufacturer/distributor*⁵ (M/D) point of view when he is deciding on how much stock to keep at each

location? He has two parameters in mind:

1. How much to keep upstream (closer to the manufacturer) in the supply chain.
2. How much to keep downstream (closer to the consumer) in the supply chain.

The natural tendency is to keep the stock as close to the consumers as possible. If a product is not at the consumption point, then there is a (much) smaller chance the item will be sold. Immediate consumption is the name of the game. Therefore, it is only logical that the M/D should keep most of the stock as close to the consumer as possible—as far downstream as he can manage—usually at the retail level. [Figure 11-1](#) shows how the inventories are distributed across a typical traditional supply chain. Most of the stock is located at the end of the chain (the shops) and little at the hub (the *plant/central warehouse* [PWH/CWH]).

The traditional supply chain displays a push behavior: pushing the products downstream toward the retailer (shop) in hopes of increasing its consumption. However, the push behavior requires a good forecasting model in order to predict what, where, and when specific stocks will be needed at a specific stock location (shop). We must have the right item (what) at the specific location (where) at the right time (when).

Why Is It Impossible to Find a Good Forecasting Model?

The advanced forecasting modules existing today try to model the demand and create a good answer to the availability puzzle: *What* product to hold at which place (*where*) and *when*. Notice that this puzzle has three questions: what, where, and when. To be a good forecast of

demand, forecasting has to answer each of these questions. The forecasting mechanism, no matter how good it is, cannot really predict what the demand would be like.

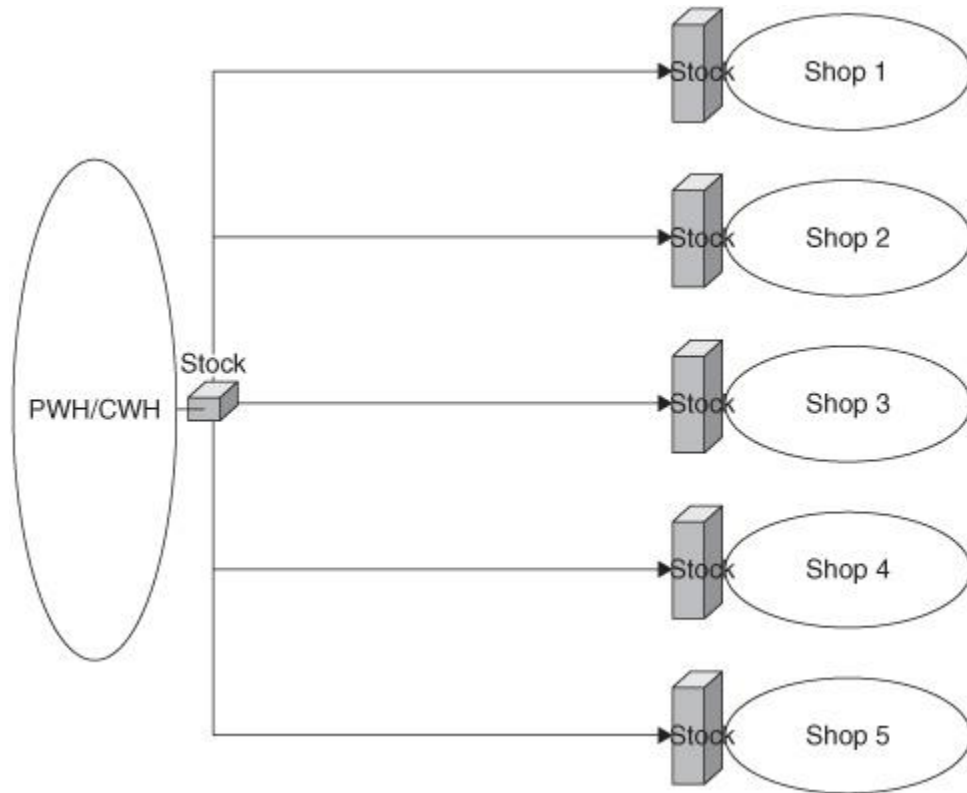


FIGURE 11-1 A typical push supply chain.

With respect to forecasting, one must consider some fallacies regarding statistics. These fallacies and a discussion of each are provided in the following sections.

1. The fallacy of disaggregation.
2. The fallacy of the mean.
3. The fallacy of the variance.
4. The fallacy of sudden changes.

The Fallacy of Disaggregation

The first fallacy is that aggregation or disaggregation has no impact on variation. The fact is that the more disaggregated the data is, the higher the variation is of those data elements. In our distribution environment, for the question of “How much demand for this product?” the answer for the M/D location is very accurate with low variability but the answer to this same question for a specific retail location is quite inaccurate with high variability. This phenomenon stems from the fact that fluctuations average out on the aggregated events (assuming they are independent events). If we predict the sales at 100 different locations, we might get an answer that sales in an average location will range from 10 to 25 units a day. If we ask the same question on the overall quantity that we need to manufacture, we will get a much narrower range as an answer—probably something ranging from 1650 to 1850. If we would just take the lows (10) and highs (25) of each of the 100 consumption points and aggregate them, we will get a much worse answer—from 1000 to 2500. This point is demonstrated in [Fig. 11-2](#). Note the high variation at the retailer versus the lower variation at the M/D warehouse. The rule then becomes the higher the aggregation, the better the forecast.

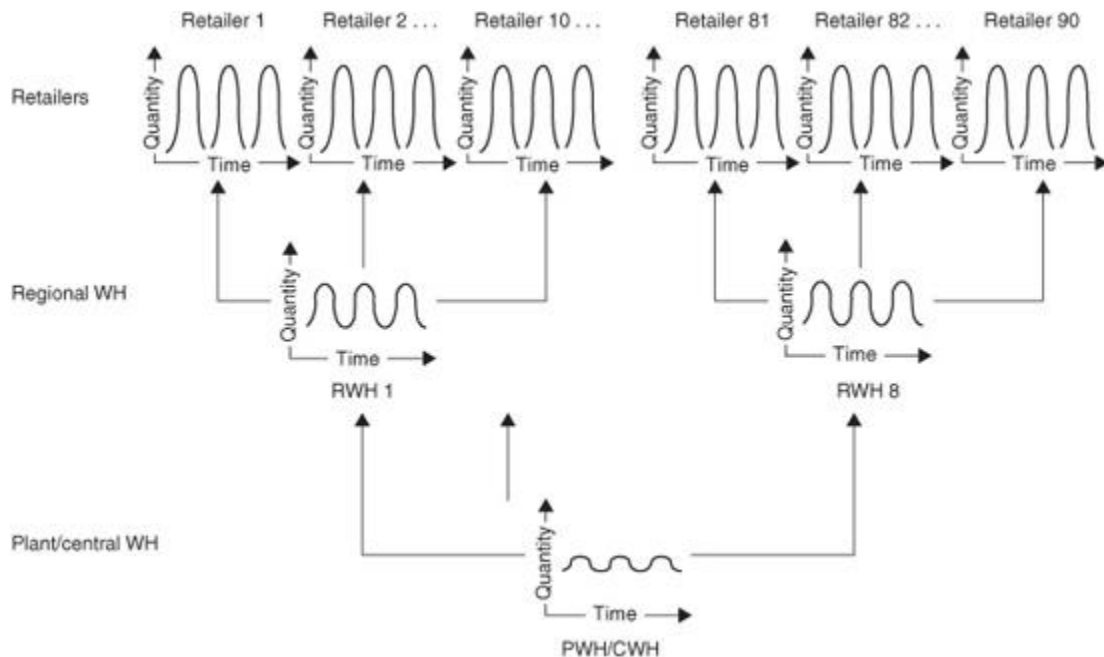


FIGURE 11-2 The mathematical effect of aggregation.

The Fallacy of the Mean

The second phenomenon relates to the wrong interpretation of data—people using statistics must have a good enough understanding of the mathematical logic that stands behind the forecast. Huge mistakes are made daily in almost every organization because of a lack of understanding of statistics. For example, the average demand in the previous example is 17.5 (assuming a normal distribution and a high and low of 10 and 25). Suppose that we stocked 17.5 units at each retail location. Do you think we would sell 1750 units? Never! There are stores that would have demand less than 17.5 units a day and we would have excess inventory (not sold) in these stores. There are other stores where we stocked 17.5 units and the demand was greater than that amount. We can only sell the 17.5 units we have that day. Therefore, overall we would sell far less than the 1750 aggregate demand. A clever man not experienced in

statistics might deduce from this example that the consumption will be between 1650 and 1850 for all consumption points, that each consumption point will have a consumption between 16.5 and 18.5, keeping 19 units for each location, and running out of stock in a fairly large number of them, while others will be left with a lot of stock they can't sell. The fact that we got an aggregated range does not mean that it can be applied to the points that make up this sum. Forecasting algorithms are getting more and more complex (software companies need to justify to the client that the new version will bring "better" results this time). One basic fact related to this complexity is important: *The more sophisticated the algorithm, the more sophisticated the end user has to be in order to use the forecast correctly.*

The Fallacy of the Variance

A related fallacy involves the understanding of variance. Most forecasting algorithms present the data as an average demand and if one really insists, then the standard deviation is given. The number of people who really understand the meaning of a standard deviation is very limited because this is a mathematical object that does not have any intuitive translation to real-life scenarios. Try to ask a salesperson not just how much he is going to sell, but also what the standard deviation is. This again calls for very sophisticated people interpreting the forecasting results in order to get some benefits from the forecast.

Suppose the salesperson estimates the average consumption of a product at a specific retail location as 17.5 with a standard deviation of 2. How much inventory should be kept at this site? If you stock exactly 17.5 units (assuming that is possible), then you would think you had a 50 percent customer service level. Recall the problem with

means stated previously. However, suppose you wanted to satisfy 95 percent⁶ of the customers requesting this product. How much should you stock? The answer is provided by the following calculations: $17.5 + 1.645(2) = 20.8$ units. Stocking only 2 units above the mean (19.5) would provide a customer service level of approximately 86 percent. The critical point is, *few people conceptually can estimate a standard deviation and determine its impact on sales without a computer.*

The Fallacy of Sudden Changes

Many forecasting methods⁷ can track changes in demand, but the more sudden the change the worst the forecast will be. An example follows. A very enthusiastic article in a paper has just appeared that suddenly changes the consumption pattern in the whole region. Suppose that the article summarizes a breakthrough study in cancer prevention and stated that if a person drinks one glass of cranberry juice a day, then this product and quantity will prevent cancer.⁸ What would happen to the demand for cranberry juice? On the other hand, suppose that a television report stated that the botulism epidemic currently spreading in our region is caused by peanut products and products containing any peanut derivative from a very large manufacturing plant in the region. What would happen instantly to the demand for these products? In today's dynamic market, such events happen frequently.

These fallacies severely affect the forecast of a single SKU (what item, where located, and when in time) and therefore provide a very poor base for determining the required stock level of that SKU. It is clear that another approach (instead of a better forecast) is needed in order to make this stocking decision.

The TOC Way—The Distribution/Replenishment Solution

The *Theory of Constraints* (TOC) analyzes the impact of the supply together with the demand to compute the right level of stocks throughout the supply chain, with the emphasis on the supply side. In an extreme case, where it is possible to respond instantly to demand, there is no need to rely on a forecast at all. While this situation is, of course, unattainable in almost all business environments, a step in this direction should be considered. In the case of keeping the right amount of stock in the supply chain, the TOC objective in responding to the three questions (what, where, and when) is to have very good availability of the items at all the consumption points (the end users). This objective is limited by the availability of cash and space, which means that it is impossible to keep high stocks of all items at all locations, even when obsolescence is not an issue. Not only that, but also as will be explained later in this chapter, keeping too high stocks of low demand SKUs will lower the total sales overall.

The *TOC distribution/replenishment solution* in the *TOCICO Dictionary* (Sullivan et al., 2007, 17) is defined as

“(a) pull distribution method that involves setting stock buffer sizes and then monitoring and replenishing inventory within a supply chain based on the actual consumption of the end user, rather than a forecast. Each link in the supply chain holds the maximum expected demand within the average replenishment time, factored by the level of unreliability in replenishment time. Each link generally receives what was shipped or sold, though this amount is adjusted up

or down when buffer management detects changes in the demand pattern.” (© TOCICO 2007, used by permission, all rights reserved.)

I will elaborate on this definition.

To respond to the three questions (what, where, when), the TOC distribution/replenishment solution is based on constant renewal of the consumed stocks from strategically placed stock buffers. The solution is comprised of six steps:

1. Aggregate stock at the highest level in the supply chain: the PWH/CWH.
2. Determine stock buffer sizes for all chain locations based on demand, supply, and replenishment lead time.
3. Increase the frequency of replenishment.
4. Manage the flow of inventories using buffers and buffer penetration.
5. Use Dynamic Buffer Management (DBM).
6. Set manufacturing priorities according to urgency in the PWH stock buffers.

Each step is discussed in the following sections.

Aggregate Stock at the Highest Level in the Supply Chain: The Plant/Central Warehouse (PWH/CWH)

The first step of the proposed TOC solution is to keep larger buffer stocks at the divergent point—where the stocks can be used to serve many different destinations—and use a pull replenishment mechanism triggered by sales at the end of the chain—the consumption point. This method

guarantees we keep the lowest stock level possible to support the demand (what, when, where) of the various consumption points (the shops).

In order to have the product available at different locations, it is recommended to aggregate the inventory at the supplying source and, when necessary, build a PWH/CWH for that purpose. When the organization is a manufacturer, the entity is called *plant warehouse* (PWH) as this is the finished goods warehouse⁹ of the plant. When the organization is a distributor, the entity is called a *central warehouse* (CWH) and is the distribution hub.

We keep most of the stock (see [Fig. 11-3](#)) at the PWH/CWH by setting the buffer stock size high. According to the principles of statistics, this aggregation of inventory guarantees a more stable and responsive system than a system of keeping large inventories at the different consumption points (shops). In the TOC solution, the amount of stock (buffer stock size) at the consumption point is very low for each SKU. When a given consumption point sells a unit, the consumed unit will be replenished as soon as possible from the PWH/CWH.

When the transportation time from the PWH/CWH to the consumption points is relatively long, a *regional warehouse* (RWH) might be needed between the PWH/CWH and the consumption points to reduce lead times. This is the case in most global supply chains and large companies where customer responsiveness is crucial to sales. An RWH pulls inventory from the PWH/CWH and ships it to the consumption points it is serving. This is just an extension of the TOC solution and all other assumptions and considerations remain the same; the idea is still to pull from the PWH/CWH based only on consumption from the consumer.

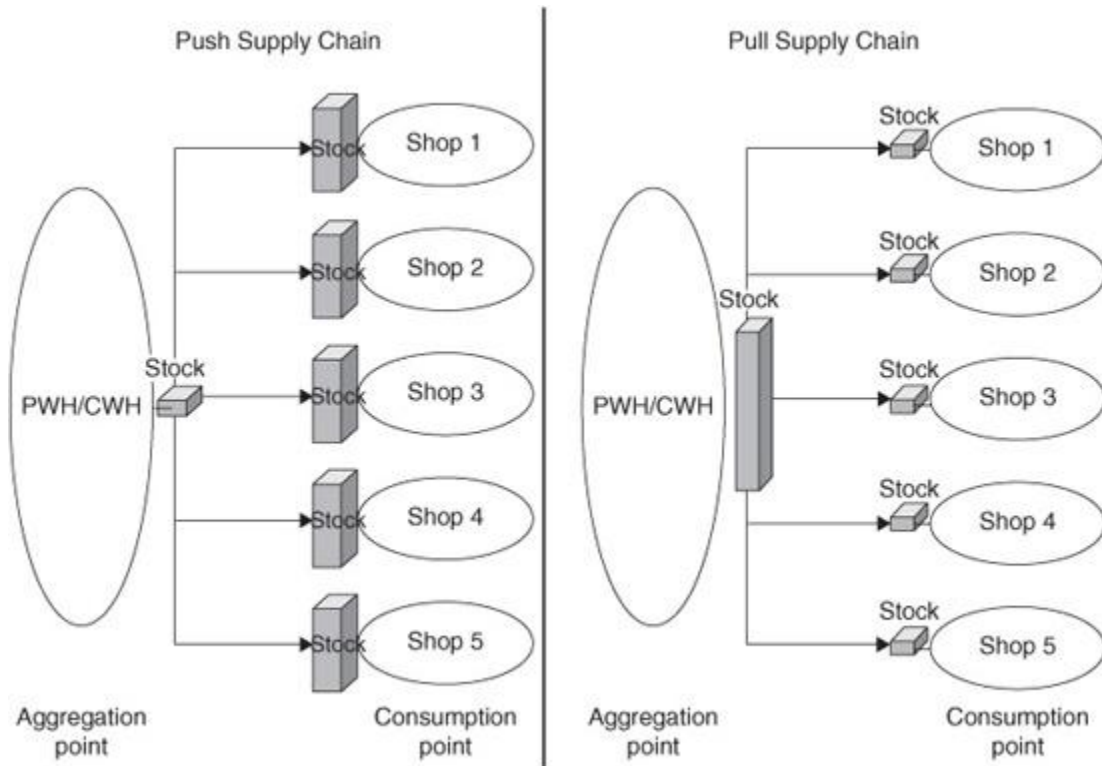


FIGURE 11-3 The push versus pull distribution supply chain model.

Determine Stock Buffer Sizes for All Chain Locations Based on Demand, Supply, and Replenishment Lead Time

The stock buffer size is the maximum amount or quantity of inventory of an item held at a location in the supply chain to protect Throughput (T). The stock buffer size (limit) is dependent upon two different factors:

1. Demand rate—*demand* is the need for an item while the *demand rate* represents the amount demanded per time period (day, week, month, etc.).

2. Supply responsiveness¹⁰—how quickly the consumed units can be replenished. The main factor here is the *TOC replenishment (lead) time (RLT)*, which is defined in the *TOCICO Dictionary* (Sullivan et al., 2007, 41) as “(t)he time it takes from when a product is sold until a replacement is available at the point of sale/use.” (© TOCICO 2007, used by permission, all rights reserved.).

A significant difference exists in the definitions of RLT in TOC and traditional push systems, and this difference and its impact should be noted before proceeding. In the *APICS Dictionary* (Blackstone, 2008, 117), the traditional definition of *replenishment lead time (RLT)* is “(t) he total period of time that elapses from the moment it is determined that a product should be reordered until the product is back on the shelf available for use.” © APICS 2008, used by permission, all rights reserved. TOC Defines this period to start from the moment the unit is consumed and not from the moment it is determined to be reordered. Looking closely at these two definitions, you notice that TOC replenishes when an item is sold or consumed versus traditional push replenishes when the quantity remaining in inventory is reduced to the reorder point (either the reorder point in the economic order quantity model or the minimum level in the min-max inventory system). This difference is significant!¹¹

Similar to the traditional RLT, the TOC RLT is comprised of three components:

1. *Order lead time*—the time it takes from the moment a unit is consumed until an order is issued to replenish it.
2. *Production lead time*—the time it takes the manufacturer/supplier from the moment he issues the order until he finishes producing it and puts it in inventory or ships it.

3. *Transportation lead time—the time it takes to actually ship the finished product from the supplying point to the stocking location.*

For example, take a regular person in his private life managing his refrigerator content. Once a week (on Monday morning), he calls the grocery store to send him two bottles of milk and some vegetables. The grocery store takes two hours to prepare the order and then another hour to send it. The order lead time in this example is one week (a whole week can pass before consumption and replenishment). The production lead time is two hours, and the transportation lead time is one hour.

[Figure 11-4](#) depicts the *traditional saw tooth diagram*. The *APICS Dictionary* (Blackstone, 2008, 122) defines this as “(a) *quantity-versus-time graphic representation of the order point/order quantity inventory system showing inventory being received and then used up and reordered.*” (© APICS 2008, used by permission, all rights reserved.) This diagram represents the reorder point/economic order quantity ROP/EOQ model and the similar min-max inventory model;¹² these models are the standard inventory models taught in many schools for managing stock levels. [Figure 11-4](#) also shows the replenishment lead time components (note that in the case of a distributor the production lead time is zero and therefore the PLT in the figure contains only the transportation lead time).

If the RLT can be reduced, then numerous desirable effects materialize:

- The amount of stock to cover demand during lead time can also be reduced.
- The amount of safety stock (to cover for uncertainty) associated with this shorter lead time is reduced.

- The forecast for new products is for a shorter time interval; hence, it is more accurate.¹³
- The responsiveness to actual demand is increased.

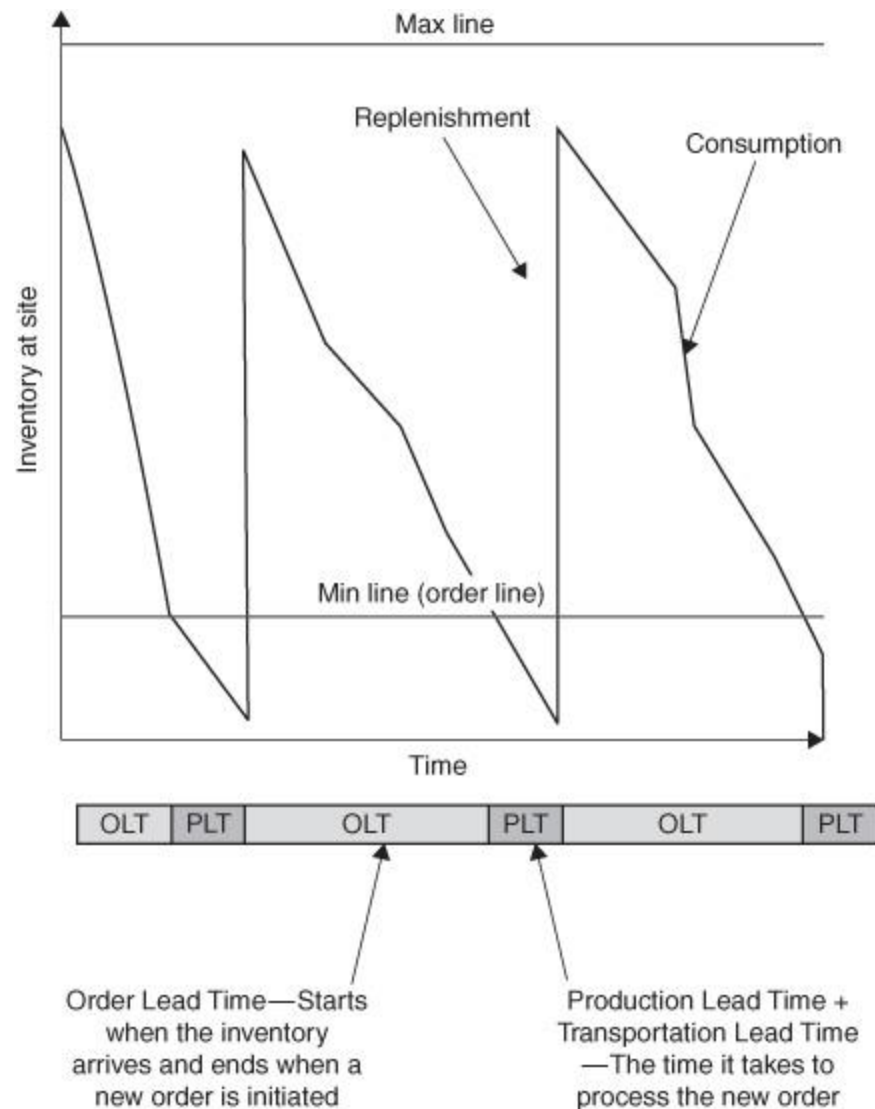


FIGURE 11-4 A typical sawtooth diagram of ROP/EOQ or min-max at the retailer or central warehouse.

These benefits (or desirable effects) make it worthwhile for you to study RLT. Try to apply these general guidelines for each component of lead time in your own supply chain.

- Order lead time—if possible, cut the order lead time to zero. For example, if you replenish each consumption point daily based on the previous day's consumption, then the maximum buffer size at the CWH for each SKU should be a few days' demand for downstream consumption points. Note that if you have an adequate CWH buffer, this becomes a reality. The only reason not to cut the order lead time to zero is if the Operating Expense (OE) goes up; this topic is discussed later. The magnitude of cutting the order lead time is demonstrated in [Fig. 11-4](#), where it is evident that just by cutting it to (almost) zero more than half of the replenishment time is saved.
- Production lead time—Simplified Drum-Buffer-Rope¹⁴ should be implemented and the priority of the manufactured parts should be tied to the actual buffer stock level at the PWH (this topic is discussed later). The PWH inventory buffer decouples production from distribution. S-DBR significantly reduces production lead time because of Buffer Management (BM) and the use of stock buffers on the shop floor to respond rapidly to the next demands not covered by finished stock in the PWH buffer.
- Transportation lead time—try to look for faster alternatives for transportation; for example, reduce the shipping interval by using trains or ships daily instead of weekly or using air shipments for some parts. Finding closer suppliers for raw materials (RMs) or purchased parts is also a possibility in many cases. Usually, this is the part of the RLT that one can do the least about, so every possibility needs to be checked. A simple test should be conducted, comparing on the one hand the extra cost of operating by faster means of transportation and on the other hand the cost saved by

keeping less inventory and the extra T generated by not having shortages. In some industries (such as the fashion industry), such a simple calculation shows it is very beneficial to gain T by using an expedited means of transportation. For example, Item A is a fashion item that is sold at a T of 80 percent from the selling price. Totally variable costs¹⁵ (TVCs) are 15 percent for RMs and 5 percent for transportation by sea. Shipping by sea has a lead time of three months. Shipping by air has a lead time of two weeks; air costs double that of sea shipment. Therefore, the T for air shipment is 75 percent. In the example, it is quite clear that shipping by air is preferable—losing a sale of one unit because of a shortage compensates for selling 15 items for a lower margin, not even counting higher inventory investment and carrying costs which are much higher when using sea shipment.

Increase the Frequency of Replenishment

When applying the TOC distribution/replenishment solution, some factors are relevant when determining the frequency of delivery. The traditional purchasing practice for managing within a supply chain encourages purchasing in large quantities. The main reasons are as follows:

1. Time and effort are required in listing all available inventories and issuing frequent orders even for a small quantity. Economies of scale exist in processing a large order versus several small ones for the buyer. However, the extra cost of managing small quantities is usually quite small and involves, at a maximum, hiring some low-salary staff to help. Sometimes a pick,

pack, and ship area is required to respond to small-quantity orders.

2. Some items can only be shipped in bulk because of transportation issues: fragile items sometimes can be better protected if shipped in a whole container; small items are stacked in boxes; and for sea and semi-trailer shipments, the minimum transportation volume is by container, making it economically beneficial to fill the container. Economies of scale exist in shipping a large order versus several small ones. New packaging might be required for the shipped items; instead of shipping a case of 48 of the same item, a mixed case of 6 units of 8 different items might be more useful. Sometimes, it is possible to use half-size containers instead of the full-size containers, so these problems can be dealt with as well.
3. Frequently, a volume discount is offered to purchase a large quantity of the same item. Additionally, a discount is given to a purchase order above a certain dollar amount. Free handling and transportation might also be given as an incentive for a larger purchase. Economies of scale exist in processing a large order versus several small ones for the seller. These discounts might be negotiated to be offered for large dollar quantities ordered over a year's period. In this way, one can order frequently and still enjoy the discount. Based on Cost World thinking (a focus on saving money everywhere), the additional shipping cost one might incur in increasing the frequency of shipments is seen as a big deterrent by most supply chain links. However, this cost is dwarfed by the increased T.

TOC takes a very different perspective, that of Throughput World thinking (a focus on making money now and in the

future), in determining the direction and frequency of replenishment. It focuses on the additional T and return on inventory investment. There is a tradeoff between the additional cost one might invest in raising the frequency of shipments and the cost of having lower availability—by making the frequency of delivery higher, a better availability is created whereas the cost of shipments increases. By making the frequency lower, one will have to pay with either lower availability or with much higher inventory levels kept at the consumption points in order to cover for variations in demand. Note, in many cases the frequent transportation will not cost more than the current large-batch transportation. While transportation costs may go up, inventory investment decreases significantly. This frees up cash that can be used to purchase product variety from the same supplier. For example, instead of having large quantities of four products from one supplier, one can invest in having smaller quantities of 10 products from the same vendor. These different products each represent opportunities for sales. In the traditional approach, the shop has four opportunities to sell to a customer; in the TOC approach, the shop has 10 opportunities to sell to the same customer. In most cases, the additional revenue produced will dwarf the extra cost. Using *Throughput Accounting* (TA) (classifying accounting numbers into T, Inventory [I], and OEs) to get estimates of the impact of increasing replenishment frequency using mixed orders (replenishing all stock buffers with each order) is an easy calculation and a profitable exercise.

For example, a manufacturer owns a fleet of cars to distribute its goods. He currently makes weekly shipments to the end points. In this example, moving to a frequency of once per day instead of once per week will generate the following:

- Increase in shipment costs—since he owns the cars, only the TVCs are added, meaning the cost of fuel and maybe hiring more drivers to fill the shifts.
- Decrease in inventory costs—instead of having a weeks' worth of inventory covering for extreme cases, he will move to keeping a daily amount of inventory. Inventory costs are effectively down by 80 percent, while the chance of running out of stock is decreased.

Manage the Flow of Inventories Using Buffers and Buffer Penetration

The TOC logic is to define the required safety and constantly monitor how the safety is being used. This safety is called a *buffer*. In a distribution environment, the quantity of an SKU we would like to keep at the stock locations (including the PWH/CWH and RWHs) is defined as *stock buffer size*. The buffer size or limit for this SKU depends on the three questions of what, where, and when to ensure high availability to support T and low inventory investment with low associated OE. For example, if the stock buffer size is 100 units for a given SKU and 40 units are currently on hand, then we expect that 60 units are on order or need to be placed on order to the supplying location. If those 60 units are not on order or on the way, a replenishment order of 60 units should be issued immediately. Note that each SKU represents an item at a location; therefore, each SKU stock buffer size may be and probably is different.

Buffer penetration is defined as *the number of missing units from the buffer divided by the stock buffer size expressed as a percentage*. The number of units missing from the buffer is the stock buffer size minus what is on hand and already ordered. For the previous example, the

buffer penetration for the stock at this site is 60 percent $((100 - 40)/100)$. The buffer size is divided into three equal zones.¹⁶ The buffer penetration sets the color of the buffer according to the different zones:

- Less than 33 percent buffer penetration: Green
- Between 33 and 67 percent buffer penetration: Yellow
- Between 67 and 100 percent buffer penetration: Red
- 100 percent buffer penetration (being stocked out): Black

The buffer penetration color gives an indication of the urgency of replenishing this stock.¹⁷

- Green—the inventory at the consumption point is high, providing more than enough protection for now. Action required: ORDER a replenishment amount (in case of replenishing from a plant, prioritize depending on whether there is enough capacity to produce this order versus more urgent orders).
- Yellow—the inventory at the consumption point is adequate. There is a need to order more units from the upstream supply chain. Action required: ORDER the replenishment amount (in the case of replenishing from a plant, order even if lacking in capacity as otherwise it might be too late. The capacity problem will be dealt with on the floor if it exists).
- Red—the inventory at the consumption point is at risk of stocking out. Units in transport/manufacturing (depending on the entity that is in charge of replenishing that stock) should be considered for expediting efforts and an urgent replenishment order must be put to the supplying source if nothing is

available on the way to the consumption point. Action required: INVESTIGATE, ORDER, AND POSSIBLY EXPEDITE.

- Black—the stock has run out at the consumption point; every hour that passes at this stage means (potential) lost sales opportunities. This situation must be resolved as soon as possible because it represents real damage, especially at the most downstream links in the supply chain (for upstream links, it means the ability to respond to replenishment and to buffer changes is diminished). Action required: EXPEDITE AND ORDER IMMEDIATELY.

[Figure 11-5](#) illustrates how the buffers are placed and how the region colors are used for prioritization. It shows the modeled network of a pull distribution system shown in [Fig. 11-3](#). The same item has different buffers, one at each location. These buffers are managed separately. The buffer in the PWH/CWH is in the size of 600 units and is currently in a buffer penetration of 20 percent (it has 480 units out of the total 600). Therefore, the priority color of this buffer is green. Likewise, in shop 1, for example, this item has a buffer of 60, out of which there are currently only 24, making the buffer penetration for this buffer 60 percent and the priority color yellow. This is how the buffers are placed and how their replenishment is prioritized at the upstream link. However, this priority is not enough, as the same buffer can have some stock in the location and some on the way.

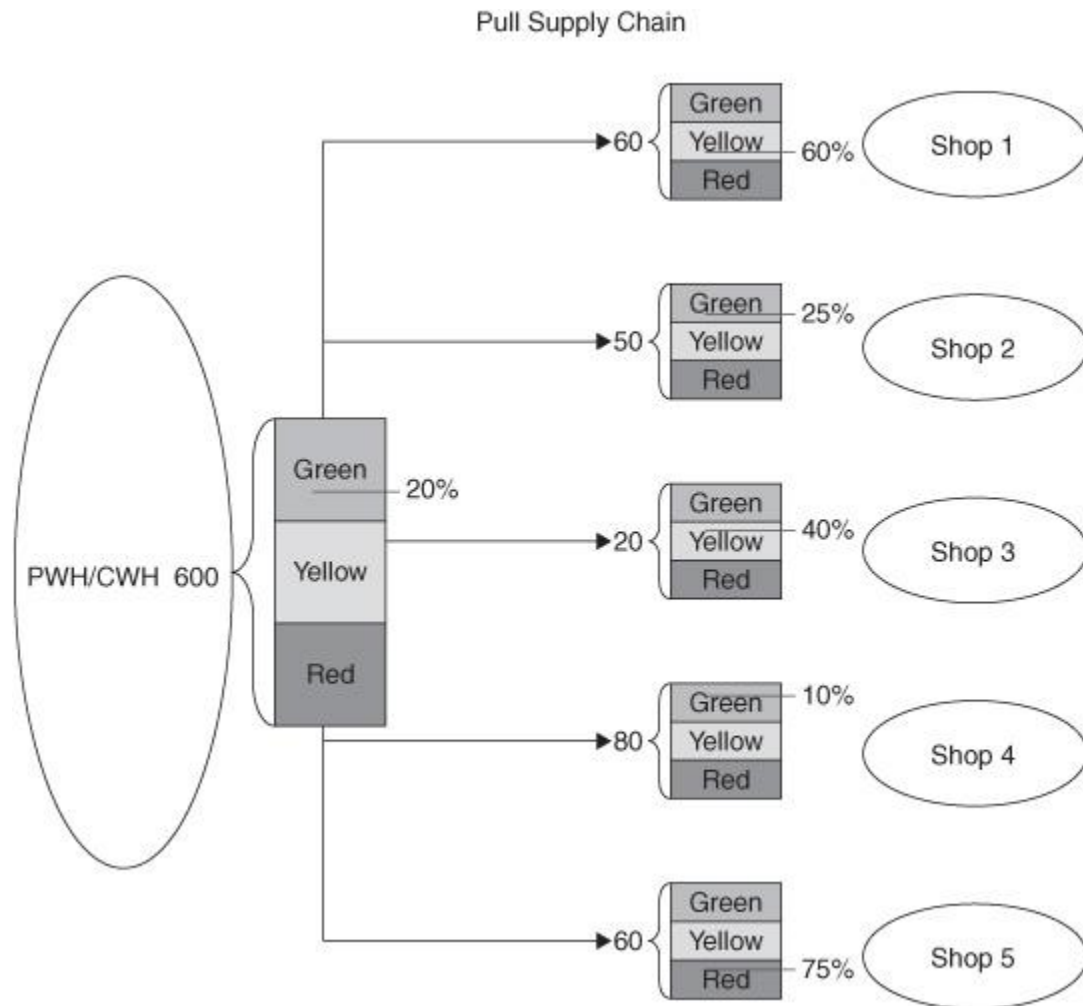


FIGURE 11-5 Item stock buffer sizes (limits) and buffer penetrations across the pull supply chain.

Several views of the same buffer are possible and of value. Inherent Simplicity¹⁸ developed the concept of the *Virtual Buffer Penetration* (VBP), which defines *the priority at any stock point as the status of the stock at the downstream links in the supply chain*. This concept is valid only until the next stocking point, meaning that the VBP for an SKU in the PWH/CWH will take into account only the physical stock at the PWH/CWH, while the VBP for a shipment will take into account the stock at previous shipments and the physical stock at the target. [Figure 11-6](#)

demonstrates this concept for managing across the supply chain.

In [Fig. 11-6](#), the retailer stock buffer size for the SKU is 100 with 25 units currently available and a shipment of 25 units on the way from the PWH/CWH to the shop. The Virtual Buffer figures appear on top of each stock on the way to the retailer. The VBP takes into account the aggregated stock of in-transit and downstream stocking points. The SKU priority is determined by the Virtual Buffer Penetration of the next downstream stock location (shown above it in [Fig. 11-6](#)). The VBP provides a very powerful tool—full visibility across the supply chain, coupled with a clear and simple priority mechanism for the various stock point decision makers involved in the supply chain. The translation of the current information for various supply chain links for this example is:

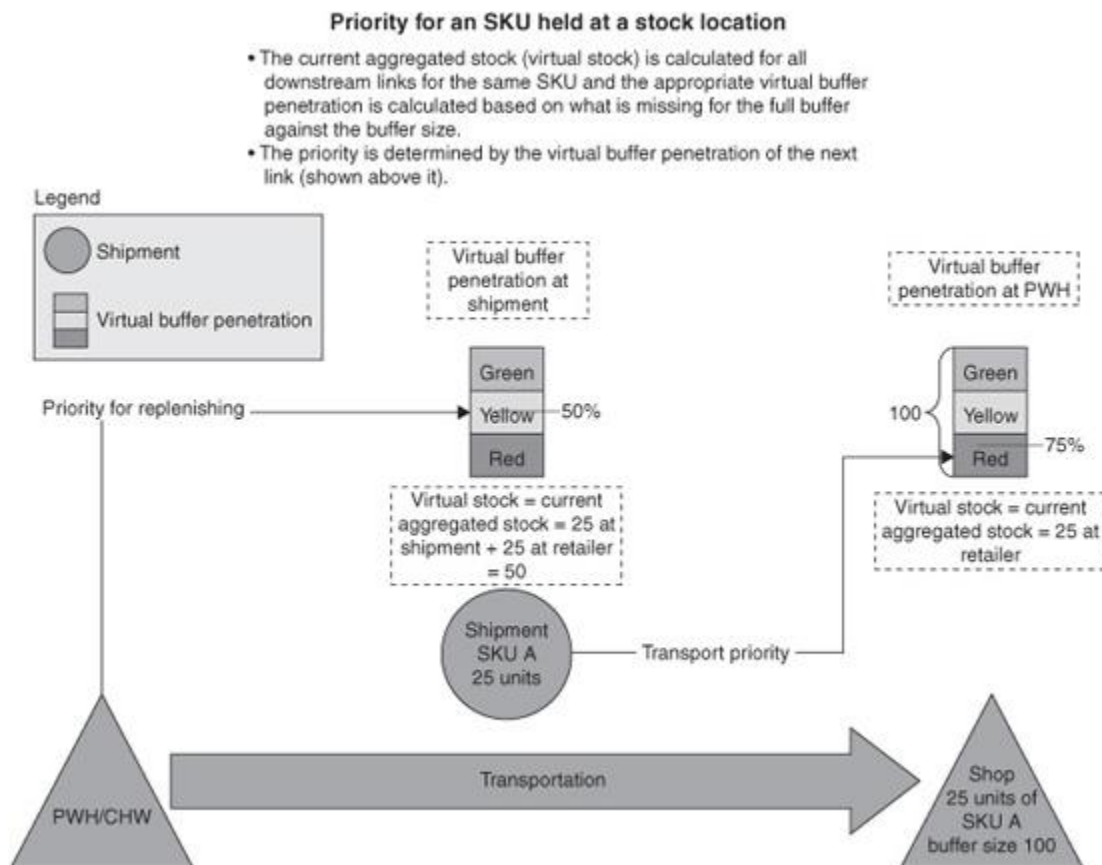


FIGURE 11-6 Virtual buffer concept applied to a shop item and in-transit shipments to this shop.

- The *warehouse manager* at the stock location (shop manager in [Fig. 11-6](#)) can see clearly that the priority of this SKU is red at 75 percent buffer penetration. The buffer size is 100 units and 25 units reside at the shop, meaning 75 are missing. The shop needs to find out how to get more stock for this SKU as soon as possible.
- The *transportation manager* can get the priority of the shipments, for example, what shipments need to be expedited. In this case, the shipment of 25 units of this SKU needs to be expedited based on a 75 percent buffer penetration (this is the same VBP as the plant warehouse manager sees). The virtual buffer for the PWH/CWH manager is computed as the shop buffer plus the transportation shipments. If the virtual buffer status was red, then the transportation manager should investigate to determine when the order will arrive at the shop. If there is some delay, he should expedite it.
- The *PWH/CWH manager* can get the replenishment priority of this SKU. This virtual buffer takes into account all stocks on the way and at the shop for this SKU. In this case, he needs to replenish 50 percent of the buffer size of this item in the PWH/CWH (50 units) and the priority of this replenishment shipment is yellow based on a 50 percent buffer penetration (the buffer size is 100 units while 25 reside on the site and 25 are on their way to the site, meaning 50 are missing).

Use Dynamic Buffer Management

TOC aims at very simple, straightforward methods so that understanding and use come easily. The concepts of stock buffer size, buffer sizing, and buffer penetration replace the need for an understanding and use of sophisticated forecasting techniques. Variations exist in reality. Therefore, Dr. Goldratt provided a mechanism to manage buffers in a dynamic environment, thus eliminating the need for these complex forecasting models. *The TOC logic dynamically measures the actual usage of the stocks and readjusts the stock buffer sizes (maximum target for replenishment) accordingly.* This method is referred to in TOC literature as *Dynamic Buffer Management (DBM)*.

By monitoring the SKU buffer penetration (i.e., item at each stock location for each product), we can identify whether the buffer size that we set for this SKU is about right. The essence of the idea is to monitor the combined impact of both the supply flowing in and the demand flowing out of the stocking point, where forecasting looks just on the demand side. The DBM approach argues that by monitoring and adjusting the buffer sizes, one can easily come to the “real” stock buffer level one needs to keep at the site in order to cover for the demand, taking into consideration the supply side (how fast one can deliver to the stock location).

The DBM mechanism is designed to alert the manager with two different warnings—one is when the buffer size is too large and the other is when the buffer size is too small.

When trying to measure whether the buffer size is too high, the indication is when actual stock of the relevant SKU compared to the target is too high for too long (e.g., staying in the green region for three consecutive replenishment periods). In other words, the stock buffer limit for that SKU should be adjusted downward, when the buffer penetration of the SKU has remained in the green zone for too long. This condition is designated as *Too Much Green (TMG)*. This means that *the stock buffer level is set too high for current*

demand. Remaining in the green zone for too long¹⁹ can be caused by the following:

- The demand rate has decreased (demand has gone down).
- The supply responsiveness has increased (the supply side has improved).
- The initial buffer size was too high.
- Demand fluctuates severely and is currently low. This is usually quite a rare statistical fluctuation. In these cases, accepting the recommendation for a stock buffer limit decrease will not reflect the reality, and therefore, soon the DBM algorithm will suggest increasing the buffer again. This condition can be caused by a downstream link offering volume discounts on a specific item or downstream links using traditional ordering models (ROP/EOQ and min-max).

The default recommendation for remaining in the green zone too long is to decrease the buffer size. The basic guideline is to decrease the buffer size by 33 percent, but this depends on several factors:

- The speed desired to lower inventories.
- The risk/importance placed on this SKU.
- The risk/importance of this stock location.

A very similar mechanism is used for determining whether the buffer size is set too low. Determine whether this SKU inventory, after replenishment, stays in the red zone. This condition is called *Too Much Red* (TMR). In other words, based on the stock buffer size *the actual stock amount*

remains after sequential replenishments in the red zone. These algorithm parameters are different from TMG, as here the risk we are trying to avoid is a stockout, while in the TMG we are trying to avoid overstocking. The basic algorithm for the TMR condition is to determine whether an SKU is in the red for several days (usually using the replenishment time as the parameter of the number of days). The more advanced algorithms also take into consideration how deep into the red the inventory at the site dropped.

The reasons for being in the TMR condition are:

- The demand rate has increased (demand has gone up).
- The supply rate has decreased (the supply side has deteriorated).
- The initial buffer size was too low.
- Demand fluctuates severely.

The guideline for relieving the TMR condition is to increase the buffer level by 33 percent. Both the definition of too long in a zone and the definitions of how much to decrease or increase the stock buffer level for each SKU are dependent on location, item, etc., and may differ across SKUs. These parameters are just good rules of thumb to establish the system.

After adjusting the buffer, the SKU needs to go through a “cooling period” in which no buffer changes are suggested until the system adjusts to the revised buffer size. This cooling period should be long enough to let the adjustment take place (the new quantities ordered should arrive at the stock location), yet short enough so that a sudden real change in the market demand will not occur without someone noticing. For the TMR, the cooling period is usually

a full replenishment time, and for the TMG, the cooling period is usually letting the inventory at the location cross-over to the green from above (since lowering the buffer size probably caused the current inventory at the site to be above the buffer size level).

Set Manufacturing Priorities According to Urgency in the PWH Stock Buffers

Many manufacturers make products to customer order. This means that each work order on the shop floor is for a specific customer with a given due date. For that environment, TOC prioritizes the production orders based on their due dates (for more details, please refer to [Chapter 9](#), which covers the make-to-order environment).

When manufacturers embrace the TOC replenishment/distribution solution, another source of demand has to be dealt with—consumption from the PWH back through the manufacturing process. For these PWH orders, the right priority for manufacturing should be set (not according to time) based on the priority of the SKU. (Recall SKU means location and we know what, where, and when given by the buffer status.) The best priority mechanism is to take the buffer penetration for the item at the PWH location (the VBP representing the physical stock at the PWH versus the buffer stock limit) as the priority for the replenishment manufacturing order, since the stock status at the PWH reflects the consumption from all downstream locations, and thus the total status of this item in the supply chain, eliminating the need for forecast. If there is more than one production order for the same SKU, the best priority mechanism is to use the VBP as illustrated in [Fig. 11-7](#).

As shown in [Fig. 11-7](#), every production order looks at the VBP of the previous production order in production (the one

that was released before it) to get its current manufacturing priority.

In this example, we see that stock in PWH for item A is 25 units versus a buffer size of 100 units. VBP is 75 percent and in the red zone. In the plant, WO1 is for 25 units, bringing the buffer to 50 units and to the middle of the yellow zone with 50 percent VBP. WO2 is for 40 units, bringing the buffer to 90 units and to almost the top of the green zone with 10 percent VBP.

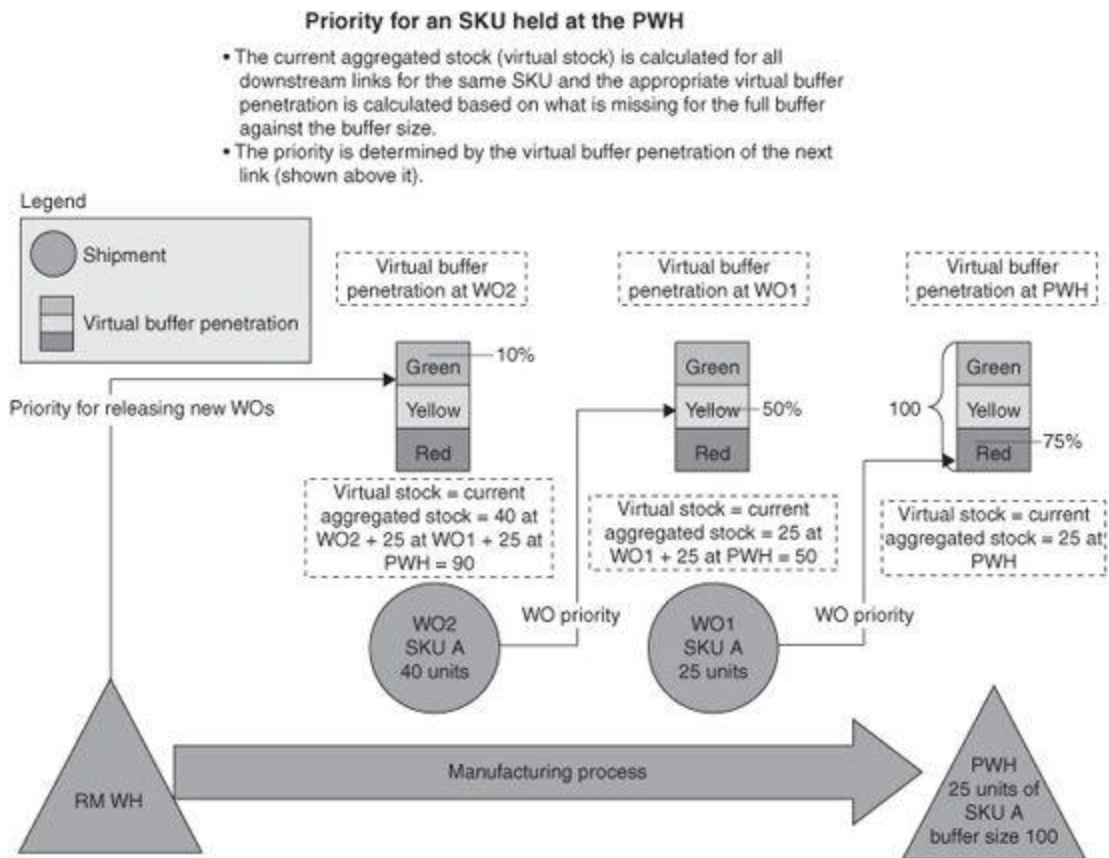


FIGURE 11-7 Virtual buffer concept applied to prioritizing work orders (WO).

This penetration measure shows that manufacturing is synchronized to the actual usage of the stock. If the stock is depleted fast, the manufacturing order will be expedited through manufacturing. Otherwise, it will follow its normal

processing sequence. Using this VBP concept provides a holistic system measure that fully aligns and synchronizes the chain links with the goal of the system—to be responsive to the actual consumption of stocks by the consumer across the chain.

Why Does a Pull Supply Chain Work Better?

Let us look at the shop and the different entities operating in this environment.

We can categorize the sale of items in the shop as three different types²⁰:

1. *Cheetah* items—these items are sold very fast relative to their stock level, enabling the retailer to reach high inventory turns²¹ (*if managed correctly*).
2. *Regular running items*—the items that do not fit the previous categories. These items generally exhibit moderate turns.
3. *Elephant items*—these items are slow movers; the retailer just can't get rid of them. These items are traditionally low inventory turn items.

What is bound to happen with the fast running items?

When items are cheetahs, by definition the market demand is high for them relative to the amount of inventory we keep for them. Regular running items (and new products) that turn out to be cheetahs are the ones most likely to be sold out. If one goes to a retailer and asks how many shortages he experiences, the most likely answer would be very few, maybe 2 to 3 percent. Misconception abounds here. If the situation and question instead were, we stand outside your store and ask people whether they found what they were looking for; in how many cases will we get

an answer of “No” even though you are supposed to carry what they are looking for? The most probable answer regarding shortages would be 10 to 15 percent. This (subjective) finding suggests that the level of shortages experienced in shops is much higher than what the retailers think.²² If the typical shopping pattern of consumers is that of purchasing more than one item at a time, then the real impact of the shortages is 10-fold. How many times have you decided not to purchase an item because the retailer was missing one or two other key items you needed? You then put the items back (hopefully) and go to another retailer hoping they have all of the items. What is the chance, when having only 15 percent shortages, for a customer to find all 20 items he wants for a home improvement project in the shop? The answer is less than 4 percent, which equals $(.85)^{20}$ as every single item has an 85 percent chance of being there but all 20 must be there at the same time for the purchase to be considered successful. These shortages affect the buying patterns of almost every customer.

A very interesting factor comes into play when analyzing those missing items: The 10 to 15 percent of items, the stockouts, are primarily the cheetahs! Hindsight being 20-20, if the retailer had known these items were cheetahs then he would have stocked a lot more. Therefore, lost sales he experiences is far more than the 10 to 15 percent that he might actually admit to! This is true especially in the fashion business. The retailers buy goods once at the beginning of the season for the whole season. Therefore, the fastest selling items (the cheetahs), which were impossible to predict a priori, once stocked out will be missing for the remainder of the season. For example, an item that sells so fast that all the inventory is consumed in two weeks in an eight-week season has lost sales of three times that which was initially purchased.

The elephant items represent the other side of the coin. These items are not sold as envisioned when the retailer bought them, otherwise he would have avoided them. The phenomenon that happens here is absurd—the retailer invests tremendous efforts to sell these elephant items and blocks his best display space with these items at the expense of the other items in the shop. This behavior, while expected from the psychological side, is counterintuitive in the business sense. Huge efforts that will be invested by the shopkeeper to sell the elephant items could have yielded much higher revenues from the cheetah items.

This phenomenon sometimes dwarfs the effect directed toward managing the shortages in the cheetah items!

Some industries have gone so far as adopting phrasings to hide the fact they are operating in a counterintuitive way in their desperation to solve these problems. They glorify the stockouts of cheetah items (according to TOC thinking, it is called lost sales) by calling them “sold out”! They then simply ignore the fact that the elephant items are bad for business by marking them “on sale” and investing huge efforts in selling them.

In a supply chain that is based on pull distribution, these negative phenomena are cut significantly. Recall that TOC BM is based on reacting to the actual market demand and adjusting the buffer sizes accordingly. If the market demand picks up (cheetah items), the stock buffer size will be increased, creating a mechanism that allows stockouts only for very limited time periods. That means lost sales due to stockouts of cheetah items are minimal. What of elephants? In the TOC distribution/replenishment solution, lower inventories of all items are kept, and the quantities are further decreased when consumption is low based on buffer penetration and dynamic buffering. Elephant items are much less of a problem as their quantities are initially low and are reduced further over time. Therefore, using pull

distribution and DBM is very effective in eliminating lost sales and overstocking.

Some of the Finer Points in Managing the TOC Distribution/Replenishment Solution

This section details some of the finer points in the implementation. Usually, those finer points come at a later stage in the implementation, after replenishing frequently and activating the DBM mechanism, but nonetheless they should be mapped at the beginning of the implementation in order to understand better and construct the implementation correctly.

Managing Product Portfolios

To differentiate between cheetah items, regular running items, and elephant items, a simple criterion exists—the inventory turns.²³ Our interest is the amount of sales of a specific item at a specific location (an SKU) relative to the inventory level of that SKU.

However, it is not enough to know the quantity in which items are sold, it is important to know their financial value as well. Just knowing from the items which are the cheetah items and which are the elephant items will not help much in driving any operational decision to improve profitability. It is important to know the magnitude in financial terms.

The goal of setting such criteria is obviously relevant when the shop owner needs to choose which items to stock. However, where a large number of items are available and the ability of each stock location to keep a large amount of SKUs is limited either by cash or by floor/shelf space, the decision is crucial. Just taking into account the inventory turns is not enough because some items are sold at such a

low margin that even if they are cheetah items they contribute little to the bottom line. In addition, a certain item can be sold only once every year (an obvious elephant item), but the margin is so high relevant to the inventory investment that it is a great item to have. For the M/D, a measurement like that can be used to support the decision of which products he should eliminate from the supply chain offering. Is there a good measure for making this decision?

The best measurement for comparing which items to stock and not stock is to determine how much a certain SKU is worth keeping at the stock location. The Return on Investment (ROI) for each inventory item²⁴ provides an excellent method of comparison across SKUs for the retailer. Retailers are usually limited by the amount of cash or space, so they should focus on the items that contribute the most to the bottom line.

Using TOC the question becomes, “How much Throughput (meaning margin) does one gain from this SKU over a year?”²⁵ This question can be written as $T = \text{selling price} - \text{TVC}$. To calculate the Investment, consider the following:

- The inventory kept at the stock location to cover immediate demand (the actual stock).
- In-transit inventory to refill the buffer. The inventory in-transit is also an investment in order to protect from the fluctuations in demand and to cover for regular consumption.

Taking these considerations into account, the best number to represent the Investment needed to generate the T this SKU realizes is the *buffer size*. By multiplying the buffer size by the TVC (TVC/unit of SKU) of this SKU, the real inventory investment needed to generate the annual T of this SKU is

realized. Note, one does not consider the timing of who owns the in-transit stock. You ordered it, and therefore there is an obligation to buy it. Hence, it should be part of the calculation.

Therefore, the formula is very simple. To calculate the ROI, the annual T of this SKU should be divided by the TVC per unit from this SKU multiplied by the (average) buffer size throughout the year.

$$\text{ROI} = \frac{\text{Annual T of SKU}}{(\text{TVC/unit} \times \text{buffer size for SKU})} \times 100\%$$

The ROI measurement enables differentiating between three different groups of SKUs based on financial contribution:

1. *Star items. These items represent a very high ROI for the retailer and certainly should be stocked appropriately throughout the chain to support the retailer. These are excellent candidates for placement at other retailers to see if consumers at those locations demand them as well.*
2. *Regular ROI items. These items are not in either category.*
3. *Black hole items. These items have a low or possibly negative ROI. These items are potential candidates for elimination from inventory. However, this is not conclusive, as some items (usually referred to as strategic) are necessary to have even though their margin is so low and/or the quantity sold is low, which places them in this group.*

It is obvious that there is a correlation between the cheetah items and the star items, but this is in no way a 1:1 correlation, as is clearly demonstrated by the extreme cases

discussed earlier and which will be further demonstrated by [Fig. 11-8](#).

The decision of how to set the limit between the different groups is dependent on the specific environment, but the general guidelines are taking the top 10 percent based on ROI as stars and the bottom 20 percent as black holes. Of course, a check is needed whether these items have been replenished regularly and are not in a class because of bad management. This check shows that you have poorly managed the inventory of the black holes. One approach to improving the ROI is to reduce the investment significantly in that SKU while maintaining its T. An obstacle that must be addressed in these situations is the purchasing unit of measure; the amount that must be purchased at one time needs to be reduced. Some items are packaged in cases of 12, 24, or 48. In any situation, the shop must sell the first 11 before it can sell the 12th item. It is far more productive to split a case and possibly get three units of four different items. You then have four opportunities to sell a product to a customer instead of one. With the cash generated from reducing the inventory investment in a black hole SKU, invest it in another SKU. Another possibility to treat black hole items is to change the price of some of those products, making them more profitable if sold at the higher prices. [Figure 11-8](#) shows an example of how these classifications can get different results.

[Figure 11-8a](#) lists 20 different items, each with its own selling price, TVC, volume, and the buffer size that had to be maintained in order to support the consumption. Assuming the buffer size was managed properly and replenishment was done properly, the calculation of the Inventory Turns and the ROI of each item appears in [Fig. 11-8b](#). The elephants and cheetahs in the IT classification and the stars and black holes in the ROI classifications are marked. Notice that while Item02 is marked as positive in both classifications, none of the other items matches the same

classification level in both cases. Especially note Item20 that achieves both a cheetah and a black hole classification, a contradiction that shows both classifications are different—ROI being the more logical one to use.

Name	Price	TVC	Throughput (=Price-TVC)	Units Sold	Cost of Inv. Sold (=Units SoldX TVC)	Buffer Size
Item01	100	50	50	100	5000	50
Item02	10	5	5	20	100	2
Item03	50	10	40	5	50	5
Item04	60	40	20	1000	40000	200
Item05	340	300	40	20	6000	5
Item06	20	15	5	50	750	10
Item07	20	5	15	20	100	10
Item08	1540	1500	40	50	75000	10
Item09	50	5	45	200	1000	100
Item10	30	5	25	500	2500	200
Item11	40	5	35	30	150	10
Item12	30	10	20	10	100	10
Item13	50	30	20	5	150	5
Item14	20	18	2	700	12600	300
Item15	70	20	50	20	400	20
Item16	200	190	10	500	95000	400
Item17	400	350	50	20	7000	20
Item18	100	80	20	60	4800	50
Item19	1000	950	50	150	142500	100
Item20	100	99	1	2000	198000	200

FIGURE 11-8 (a) The data for the calculation of item inventory turns, ROI, and their item classifications.

Inventory Value	T Generated	Inv. Turns	ROI	IT class.	ROI class.
(=Buffer Size* TVC)	(=Units Sold* T)	(=Cost Of Inv. Sold /Inventory Value)	(=T Generated/ Inventory Value)		
2500	5000	2	2	Normal	Regular
10	100	10	10	Cheetah	Star
50	200	1	4	Elephant	Regular
8000	20000	5	2.5	Normal	Regular
1500	800	4	0.5333333333	Normal	Regular
150	250	5	1.666666667	Normal	Regular
50	300	2	6	Normal	Regular
15000	2000	5	0.1333333333	Normal	Black Hole
500	9000	2	18	Normal	Star
1000	12500	2.5	12.5	Normal	Star
50	1050	3	21	Normal	Star
100	200	1	2	Elephant	Regular
150	100	1	0.666666667	Elephant	Regular
5400	1400	2.333333333	0.259259259	Normal	Regular
400	1000	1	2.5	Elephant	Regular
76000	5000	1.25	0.065789474	Normal	Black Hole
7000	1000	1	0.142857143	Elephant	Regular
4000	1200	1.2	0.3	Normal	Regular
95000	7500	1.5	0.078947368	Normal	Black Hole
19800	2000	10	0.101010101	Cheetah	Black Hole

FIGURE 11-8 (b) The calculation of item inventory turns, ROI, and their inventory and ROI item classifications.

Rules for Setting up Initial Buffer Sizes

The first step in moving from push distribution to pull distribution is setting up the PWH and starting to build inventories to fill the initial stock buffers.

The decision of what size the initial stock buffer should be might seem to be a very complex decision as the amount of uncertainty is huge. Fear of making an error or the wrong decision and jeopardizing the whole initiative is natural.

The answer of how big the buffer size should be is quite simple. There are not enough words in the dictionary to emphasize the difference here between being *precisely wrong* and *approximately right*. It is not exceptional to find cases in which determining the initial buffer targets took more than three months! Starting with any initial guess and adjusting the buffer size according to DBM would have

reached good enough buffer sizes much faster. Based on the parameters (demand rate and supply responsiveness), a generous stock buffer size can be determined (which is generally much lower than what is stocked now in the chain). Since the DBM mechanism will adjust the buffer sizes according to real consumption, the initial estimates are not that critical.

It is advisable to start with an initial guesstimate: taking the replenishment time from the source to the destination and multiplying it by the average daily consumption and by a factor (to cover statistical fluctuations). For the PWH/CWH, a fluctuation factor²⁶ of 1.5 is appropriate. For the selling points, a factor of 2 is appropriate, as the fluctuations are larger there. The replenishment time to use should be:

- For a production environment (PWH) taking the current quoted production lead time for this item (after implementing TOC in the manufacturing environment, the lead time will usually be cut in half). Use this lead time and remember that DBM will automatically suggest lowering or raising the stock buffer level over time.
- For a distribution environment (CWH, regional warehouse, and consumption points), transportation time plus something to account for a low weekly frequency of delivery if needed.

One must also adjust for the frequency of delivery. As the frequency increases, the buffer size is smaller; as the frequency decreases, the buffer size is larger. For example, it is obvious the buffer would be much smaller if the SKU was delivered daily versus if it was delivered every week.

Managing Seasonality in the TOC Distribution/Replenishment Model

Evidence supports the claim that the DBM is an excellent mechanism to monitor and control stock levels when changes in supply or demand are gradual²⁷ or otherwise when the changes are unpredictable.

However, sudden large changes in either supply or demand are not handled well by the DBM mechanism or by any other known mechanism. An unanticipated, sudden, and steep increase in demand or deterioration in supply can cause shortages that will lead to lost sales and a damaged reputation. Alternatively, an unanticipated, sudden, and steep decrease in demand or improvement in supply will cause excess inventories and an undesired focus on sales efforts to move slow-running items.

While some of these changes in patterns are unpredictable (and thus unanticipated), experience shows that there are known recurring predictable, sudden, and steep patterns. These situations can be moderated by recognizing where and when to use the DBM methodology. The general guidelines are as follows:

- When the changes are gradual, either predictable or unpredictable, use DBM—this should be the mostly used method. Preferably, use automatic DBM in order to avoid mistakes and enhance the focus on exceptions.
- When the changes are unpredictable and large, use DBM in order to point to the problem and use a manual decision process to define by how much the buffers should be changed when the change occurs.
- When the changes are predictable and large, use the seasonality module. The seasonality module handles known patterns for sudden changes in consumption.

Known Patterns for Sudden Changes in Consumption

Most of the time significant changes in supply or demand are predictable. Marketing and Sales people know from experience when to expect these changes in demand and their consequences on the supply. Generally, the direction of the change is well known and a gross approximation of the size of the change is possible, enabling taking measures ahead of time to deal with the change. Typical causes of changes in patterns²⁸ are divided into two groups: Pull Seasonality & Push Seasonality.

Inherent Simplicity defines the following patterns as *Pull Seasonality*, meaning the environment defines the demand pattern for the organization without the organization being able to do anything about it:

- Seasons in the year affecting the consumption of certain SKUs.
- Holidays or events geographically affecting where certain SKUs will be consumed, more or less.

Inherent Simplicity defines the following patterns as *Push Seasonality*, meaning the organization, for various reasons that should be verified, takes actions that create a peak in demand in the market. These patterns include the following:

- Promotions—very similar to holidays in nature as they are short and create a spike in demand. They are generally followed by a low period in demand.

- A known price increase—many times an organization will announce an increase in product prices becoming effective at a point in time. Customers generally stock up on the item before the price increase. They are generally followed by a low period in demand.
- Financial period end seasonality—measurements of salespeople that focus on quarterly or yearly quotas usually create a kind of seasonality in which sales go up sharply before the period ends and down sharply at the start of the next period, caused by pulling orders ahead toward the end of the period. Note: this can also be created from the budget management of the clients—toward the end of a period, they try to take advantage of all unused budgets for purchasing. They are generally followed by a low period in demand.

Two Different Changes

Each of these situations has a beginning and an end, which Inherent Simplicity²⁹ defines as *Sharp Demand Changes* (SDCs). In the previous descriptions, the beginning SDC is (usually) the event that causes an increase in demand and the end SDC signals the sudden end of the increased demand and a return to “normal” or below normal demand.

Resolving the Forecasting versus DBM Dilemma to Provide Excellent Consumption before, during, and after an SDC

SDCs present a problem in changing the buffer sizes. What would happen if DBM will continue to be used in an SDC? A possible problem is demonstrated in [Fig. 11-9](#).

The dashed line in [Fig. 11-9](#) (Inventory) represents the actual inventory at the site. The inventory is more or less

stable until the season starts. After the season starts, since there is a huge surge in demand (Sharp Demand Increase in the figure), the on-hand inventory runs out completely. The DBM mechanism almost immediately suggests adjusting the buffer by 33 percent (from 9 to 13). The stock buffer inventory stays on zero for some time (any small replenishment orders that are already in processing are consumed immediately because the demand is so high), which triggers the stock buffer level to be increased by 33 percent (DBM first buffer increase in the figure). During this period, sales are potentially lost because demand is higher than supply. When the new replenishment quantity arrives, it is still not enough to support the new demand because the demand picked up by much more than 33 percent. That causes the same phenomenon to occur—the inventory runs to zero until the new replenishment quantity arrives at the site, representing another 33 percent increase—DBM second buffer increase (from 13 to 17 in the figure). By the time that quantity arrives, the demand has already gone down and the site is left with too much inventory to support the demand. The DBM mechanism identifies that condition but by that time, it can only reduce the buffer by 33 percent, leaving it much higher than it should be in order to properly support the demand. It will eventually reach the steady-state level again, but for some time you first experience stockouts followed by carrying excess inventories. Of course, this is an extreme case, but obviously must be dealt with.

How will DBM cope?

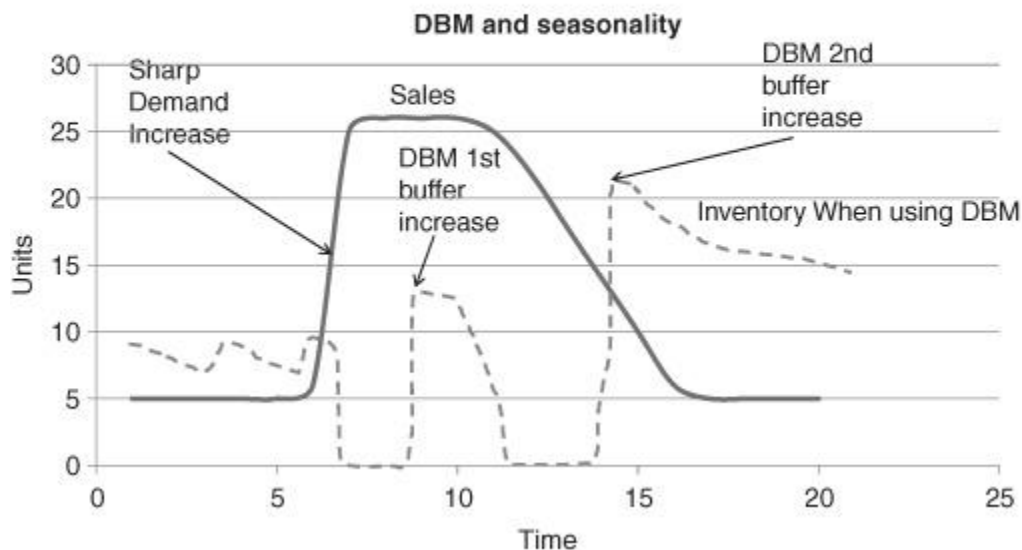


FIGURE 11-9 The problem of using DBM with an SDC. (© 2007 *Inherent Simplicity*. All rights reserved.)

It is apparent that sometimes crude forecasting must be used in order to avoid those negative effects. [30](#)

Identifying When an SDC Is Meaningful

A simple rule can be used to determine whether an SKU is exposed to certain seasonality effects. Look back on last year's consumption (and the year before, if possible). If one month's sales are more than twice the monthly average of the total sales (greater than approximately 15 percent of the whole year, say the Christmas season, for example), then this SKU should be looked at carefully to see whether the SKU is an SDC item. While DBM reacts to reality quickly, using seasonality forecasting does not (the shop must adjust orders manually). Therefore, it is important to define an SKU as seasonal only if it creates a huge difference with which DBM cannot cope. Most changes, especially when the

replenishment time is relatively short, can be easily dealt with by DBM. If the order frequency is over a day or two and the spike in demand is high and short, you should adjust the orders manually. If ordering daily with a short replenishment time, a change as high as a 50 percent increase in consumption in the course of one single replenishment time is something with which DBM can usually cope.

Handling of an SDC

For known SDCs, you need to forewarn the upstream links in enough lead time to respond. If the spike in demand timing is known (e.g., a home football game) and big (much larger than average demand), then you should give the upstream links notification to be able to respond and plan how the SDC should be handled. When an SDC is identified, it should be treated in the following manner,³¹ depending on the direction of the SDC.

For a known large SDC that marks an increase in demand (also defined as Sudden Demand Increase):

1. Stock buildup.
2. Disable the DBM (cooling period).
3. Back to normal (or sometimes even below normal).

For a known large SDC that marks a decrease in demand (also defined as Sudden Demand Decrease):

1. Stock builddown.
2. Disable the DBM (cooling period).
3. Back to normal (or sometimes even above normal).

Note that Steps 2 and 3 in both cases are the same—the same actions need to be performed in order to treat these SDCs correctly. Only the first step is different in both cases.

[Figure 11-10](#) describes a typical SDC with these steps across time, containing the management of two consecutive SDCs—one sudden increase and then a gradual decrease in demand.

Stock Buildup

In this phase, the purchase order is issued to the supplier to replenish the stock to the forecasted buffer level *after* the SDC (notice that the changed demand pattern [sales] is after the SDC and not during it). Two different environments (distributor and manufacturer) exist. Each is a little different.

- Distributor situation: If the SKU is a purchased item, and the supplier has no problems supplying the larger quantities (the required quantity is the difference between before and after the SDC), the best way to perform the buildup is a one-time order from the supplier. The order should be received a full supplier lead time before the suspected start of the SDC, with some extra time buffer to cover for the supplier's unreliability (Murphy always strikes). For example, a distributor holds a buffer of 200 units from a specific item in the CWH to manage normal consumption. The regular lead time of this supplier is two weeks. During Christmas, he knows consumption doubles. He should double the buffer (to 400) of this SKU approximately 2.5 to 3 weeks before Christmas sales increase.

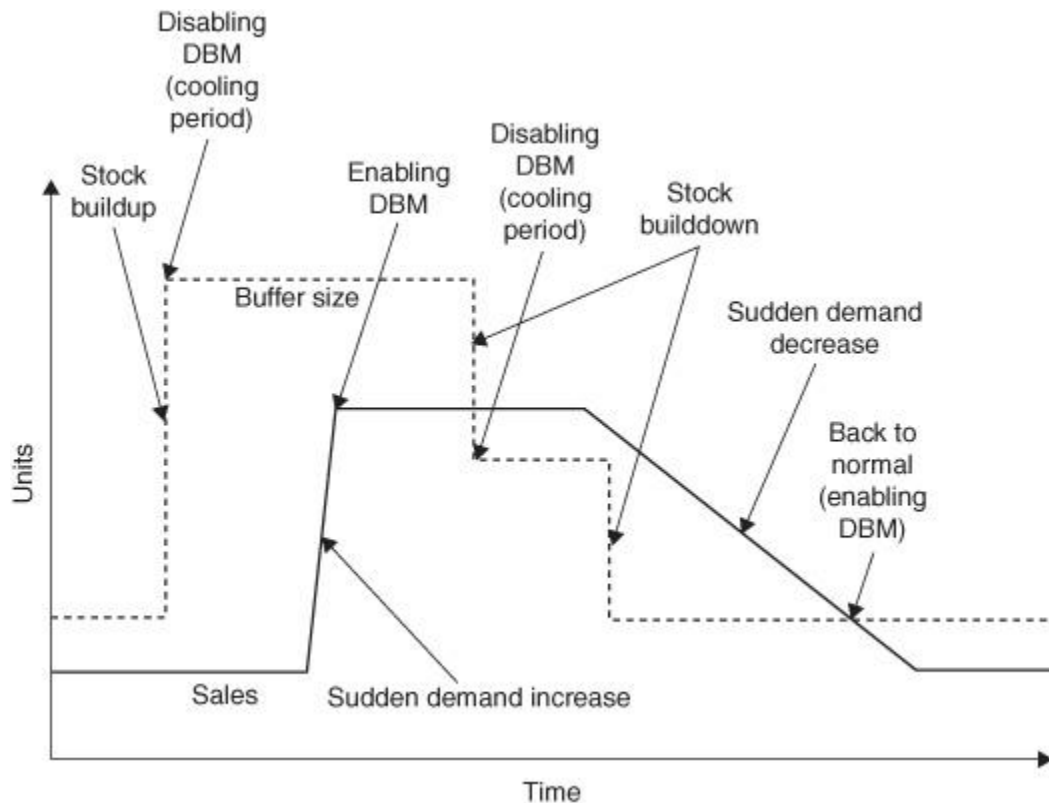


FIGURE 11-10 The Inherent Simplicity steps for managing a typical SDC. (© 2010 *Inherent Simplicity*. All rights reserved.)

- **Manufacturer situation:** If the SKU is manufactured by us, or by another manufacturing supplier that cannot support providing large quantities from his inventory, the best way to perform the buildup is to manufacture/order the missing quantity in batches, over a longer period of time, depending on our production capacity or on our supplier's ability to supply. Again, provide a time buffer by requesting delivery one receipt cycle ahead. For example, let us take the same case as above, only this time a batch of 80 units per 2 weeks is the maximum manufacturing can handle. In this case, the buffer should be adjusted three times—once

increased by 40 units approximately 7 weeks before Christmas sales start, then by 80 units 5 weeks before, and then by another 80 units 3 weeks before. Sales should be monitored to ensure that the orders are appropriate.

Disabling the DBM (Cooling Period)

After changing the buffer size to reflect the future demand, disable the DBM algorithm in the same way it is disabled during the cooling period after a DBM buffer change.³² It is important that while the changes are realized, the DBM will not start operating as the whole purpose of treating the SDCs this way is to ignore the reality because we have better knowledge of the future reality. Normal DBM activity will disrupt the proper handling of an SDC and might have very negative ramifications; hence, the need to disable it during this time.

Back to Normal

After realizing the changed buffer size at the CWH/PWH/RWHs to get ready for the different future demand, the SDC occurs within a small time frame. It is important for the chain to be very responsive in the replenishment to the consumption points and in the decisions to increase or decrease the buffers at the various consumption points according to the DBM mechanism after the SDC.

Stock Builddown

Usually, a Sudden Demand Increase³³ is accompanied by a Sudden Demand Decrease, which brings the demand for the

SKU back to “normal.” Sometimes, this situation is less problematic because the demand drops very gradually, allowing the DBM to adjust. The traditional after-Christmas sales and end-of-year sales are an example of a more gradual decline caused by the dumping of excess inventories at significantly lower prices. The point is that it is very important to try to refrain from being left with excess inventory after the Sudden Demand Decrease. Being left with large amounts of SKUs after an SDC will focus salespeople’s attention on the wrong products, will force the consumption points to offer huge discounts on the SKUs, and will occupy shelf space, otherwise much better used for the star items. It also establishes a consumer demand pattern of waiting until after Christmas to buy items.

The builddown is very similar to the buildup of inventory—it is important to decide whether the reduction of the stock in the system will be done in one step or in a few steps—depending on the steepness, the demand drops. Usually, the demand drops gradually and therefore it is best to absorb the reduction in a few steps. Important note: just as the increase in inventory was planned over a period of time, the decrease in inventory should be planned and will take time. Depending on the expected aggregated demand until the peak demand falls, it is important to set the buffer size to be decreased, stopping the replenishment well ahead of the demand decrease and in this way ensuring that one is not left with too much stock at the end of the peak. In most of the cases, it is enough to lower the stock buffer size (target) about a replenishment time before the demand is expected to start dropping, thereby stopping the replenishment of the SKU until the amount of inventory on hand goes below the (new) buffer size. Replenishing within the peak is done under stress. Suppliers and distribution channels feel the high demand and are under pressure. Therefore, it is important to reduce the pressure on some of

the items—those that we don't need until the end of the peak—maintaining the focus of everyone involved.

Implementing the TOC Distribution/Replenishment Model—How Can Software Help and Is It Really Needed?

To successfully implement the TOC methodology to manage a distribution environment, three major requirements need to be fulfilled:

1. Replenishment—replenish stock buffer size according to consumption at all locations.
2. DBM—manage the stock buffer size constantly at all stock locations and adjust it to support the consumption points' demand. It is very important, especially for environments that manage a large number of buffers, to have the software automatically manage the DBM changes.
3. Managing predictable SDCs—override the DBM mechanism during SDC buffer changes.

These requirements are not the only ones that need to be implemented, but these three are necessary conditions for success in any manufacturer/distribution implementation of the TOC principles.

Even considering only these three requirements, the conclusion must be that no distribution organization can manage based on TOC without software, unless it is a really small distribution chain (anything more than 100 buffers to manage requires some kind of software or additional personnel). The question is, what kind of software should be used?

First, define how many buffers are likely to be kept under the TOC distribution/replenishment model³⁴:

- The number of items that will be managed—this is the number of items the company currently offers the market.
- The number of stock locations in which the items will be managed—all warehouses (PWH, RWH) and client shops for the manufacturing environment and all warehouses (CWH, RWH) and client shops in which the future the SKUs will be stocked in a distributor environment.

The estimate on the number of SKUs and therefore buffers that will need to be managed stems from the multiplication of the two previous items.

In general, there are three options to choosing software:

1. Develop the needed software components within the existing ERP system used by the organization.
2. Develop the needed software components as Excel sheets external to the ERP system.
3. Purchase an external TOC distribution/replenishment solution software.

The answer to the question of which option to choose depends mainly on the operation scale.

- For any environment where less than 500 buffers are required, using internal software is a possibility (whether an Excel sheet or a development of the current IT system).

- For any environment where more than 500 buffers are required, the recommended solution is to get external software, which is fully focused on the TOC processes and decision making.
- For an environment where an ERP system is operating effectively and more than 500 buffers are required, the IT staff should read and study this chapter closely before undertaking the development and integration of the TOC distribution/replenishment solution into their existing ERP. This is far from an easy task, and is usually not recommended due to the reasons that follow. One should also recognize that the TOC distribution/replenishment solution works best where S-DBR is fully functional. If this is the chosen option, consulting should be used for the process.

The benefits of using external TOC software over developing an internal solution are the following:

1. Quality assurance—Ensuring that the internally developed software module is doing what it should is very problematic. The good TOC add-on software vendors are investing most of their efforts on checking the validity of the modules.
2. Reliability—Ensuring that now and in the future, no changes or additions are made to the modules (causing negative ramifications) by people who “think they know the philosophy and the environment.”
3. Development—The TOC body of knowledge in distribution/replenishment is currently growing rapidly. TOC consultants and software companies develop new insights continuously, and the TOC software companies invest time and efforts in order to

incorporate the latest knowledge into their software. Unless you have a highly skilled TOC expert leading or advising the company continuously and dedicated software designers developing the distribution/replenishment functionality, an internally developed system will never keep up with the developments in the field.

4. Proper know-how—Many fine details are not within the public knowledge domain. When considering companies with special needs, such as seasonal products, limited shelf-life products, fashion products, and groups of similar products or large numbers of buffers, only a TOC software company can incorporate software modules to correspond with those needs without significant investment in time and money to determine how to handle the environment and product characteristics.
5. Long development lead time—Based on our experience trying to help other companies develop and test the logic for distribution/replenishment software modules incorporating the environment and product characteristics, the time needed is significant. It usually takes at least twice the amount of time originally planned to complete the development—usually between six months to two years for full functionality.
6. The Excel problem—While Excel is an excellent tool for many applications, an Excel sheet, despite its relative ease in building and use, is especially not recommended. An Excel sheet is very easy to change. Anyone, including people without the proper knowledge of the distribution/replenishment solution, has the ability to modify it on purpose or by accident,

and therefore it cannot really be used in order to enforce the correct use of the distribution/replenishment solution. Additionally, an Excel sheet is very hard to debug. Both quality and reliability are issues in the use of Excel sheets for this application.

Testing the Solution on a Smaller Scale

The TOC distribution/replenishment solution can be tested in two forms prior to implementation. Both forms have advantages and disadvantages.

Simulation

It is possible to do a kind of “simulation” in order to show what results can be achieved prior to implementing the TOC distribution/replenishment solution for a specific environment. In the simulation, the real consumption data and stocking level figures can be simulated and benchmarked against the historical data. This same data can then be used with the TOC distribution/replenishment solution to provide a partial result for comparing against the current environment and its traditional distribution/replenishment methods. The comparison should show the impact of the changes in policies, procedures, etc., on inventory levels, investment, OE, stockouts, and service levels for this specific environment. TOC will increase availability, while reducing the total inventories held. A typical simulation result we have conducted in Inherent Simplicity is shown in [Fig. 11-11](#).

Simulation Summary Report

Properties
Set Report Properties

Export

Print

Results from 20/02/2009 to 30/05/2009

	Simulation Results	Reference Results	Improvement
Inventory			
Average Inventory	5,129	9,256	-44.59%
Closing Inventory	1,740	6,304	-72.40%
Availability			
Stock Out Days	2	48	-95.83%
Stock Out Days %	0.22%	5.39%	-95.92%
Sales			
Lost Throughput (TxD)	63	264	-91.76%
Inventory Turns (ITR)	16.17	9.03	79.20%
Return On Investment (ROI)	1.29	0.72	79.46%

FIGURE 11-11 An Inherent Simplicity simulation example. (© 2010 Inherent Simplicity. All rights reserved.)

However, this simulation has a few significant drawbacks that are very important to emphasize. The first two drawbacks are general and apply to most simulations:

1. A simulation is based on certain assumptions (such as the actual replenishment time, frequency of replenishment, etc.); one invalid assumption might cause a very different result in the simulation versus real life.
2. Human behavior cannot be simulated by the computer unless some very specific assumptions are modeled, ones that will not be simple to quantify.

The second drawback is quite large, and could cause the following six misalignments between the simulated state and what would have happened in reality had the TOC solution been implemented. The first three of the six emphasize the focus on T, on which the old environment failed to capitalize; the second three examine the impact on OE.

1. More sales are generated by the TOC solution due to elimination of item shortages in the real-life situation. Since there was no stock, a sale could not occur even though in the simulated state the stock might have existed. Additionally, recognize that these lost sales might be as high as 15 percent! (Reality will have *better* results than the simulation.)
2. More sales are generated by the TOC solution due to the change in the retailers' focus from slow- to fast-moving items. (Reality will have *better* results than the simulation.)
3. In the longer-term, more sales are generated through the TOC solution due to the improvement in the company reputation for short lead times and high due date performance. (Reality will have *better* results than the simulation.)
4. Less obsolescence (depends on the environment) is achieved by the TOC solution due to the higher inventory turns. This can be calculated roughly based on the difference in inventory turns between the simulated state and the real-life state. (Reality will have *better* results than the simulation.)
5. Frequently, higher transportation costs are generated by the TOC solution. This can be calculated based on the assumptions on frequency of replenishment, although usually other factors might affect the calculation, such as the rate of acceptance of the suppliers to switch to rapid replenishment. (Reality will have *worse* results than the simulation.)
6. In the TOC solution, cross-shipments are virtually eliminated. Expediting shipments are also reduced

significantly. (Reality will have *better* results than the simulation.)

Because of these drawbacks, the simulation is only useful in giving a general direction of the solution and for buy-in purposes because it usually underestimates the magnitude of benefits of the TOC solution.

Pilot Project

Running a pilot project on a small part of the business prior to implementing the solution across the business is a valid way to test the solution and its ramifications. For a large organization, starting the solution on a part of the system makes a lot of sense. The following points are important to note while conducting a pilot project.

- Design the pilot based on valid test parameters. The pilot and control group test sites should be selected so that the pilot and control group results are meaningful for the current set of conditions (economic, organization, product, etc.). The historical data of results should be similar for these two sites. The distribution/replenishment pilot is then implemented over a sufficient test period (generally three to six months). The test period should be long enough to eliminate the impact of the starting conditions, to get down the learning curve on how to manage with DBM, and to experience some of the difficulties of this environment. The results of the pilot are compared to both its historical results and to the control group results.
- Define in advance what is to be measured in the pilot. Generally, growth in sales, number of stockouts, length of stockouts (measure of exposure to lost sales), service

levels (impact on T), inventory levels (impact on I), expediting and overtime costs (impact on OE), lead time, and due date performance (impact on future T) are excellent measures. The results of the inventory turns or ROI (both macro measures) should be checked and compared as well.

- Equally important is to determine the decision criteria for the pilot ahead of time. What do the results have to be for the TOC distribution/replenishment to be deemed a success and warrant full implementation? When should it be abandoned? For example, if the inventories of the pilot compared to the control group and its historical data are reduced by 30 percent while availability improves (increased T) and the other measurements did not suffer deterioration, then the TOC solution will be implemented in all of the RWHs.

Other considerations are as follows:

- The most upstream point in the supply chain under the pilot organization's control (typically the PWH/CWH) should be part of the pilot—at least for the chosen SKU portfolio to support the pilot chain flow.
- It is advisable to include some downstream nodes for the same SKUs as the effect of the pull distribution solution will be higher the closer the implementation is to the actual consumers.
- If the pilot is run on a PWH, it must be possible to give higher priorities to the pilot SKUs over the control group SKUs in order to show the benefits. Otherwise, it is imperative to hold some safety stocks that will be triggered should the replenishment time pass without actual replenishment. This answers the question, "If I

have implemented the full TOC distribution/replenishment solution, will I be able to respond this rapidly to the chain needs?"

- The pilot should manage a minimum of at least 100 buffers.
- The same items should be managed in both environments.
- Most of the buffers (at least 50 percent) of the pilot need to belong to fast-running items. This is to demonstrate the difference in focus on T caused by focusing on fast versus slow movers in the two different environments. Both the pilot and the control group have the same items but certainly, the inventories will have a significant difference on the retailer focus.
- A sample of buffers can belong to slow-running SKUs to test where the best decision point would be between managing items to be held for availability and items managed to order.

Managing the TOC Buy-in Process

Managing a change in an organization is never an easy task. Implementing TOC adds some complexities, as the underlying message brought to the organization that embarks on TOC might be interpreted as, "It's so simple you should have thought of that by yourself." The TOC pull system is much simpler than traditional push systems and generates better results. However, any change is not a trivial process. It encompasses breaking old habits and this is difficult. Implementing TOC requires breaking several old habits. Therefore, it is a challenging task although the new processes are simple. Some of the changes TOC brings are as follows:

- A paradigm shift is involved. TOC challenges the most basic assumptions of traditional management: the focus on cost saving everywhere (for example, ordering big batches, moving big batches, storing big batches, and selling big batches everywhere in the chain). Therefore, ongoing training is essential to understand the impact of these cost-savings actions and how TOC treats these same decisions.
- New processes are introduced. The introduction of the PWH/CWH by itself involves several new processes (for example, shifting inventory control from make-to-stock to make-to-availability), as well as determining new methods to handle seasonality, moving to daily replenishment, etc.
- A lot of data needs to be collected, processed, and managed. The TOC pull distribution/replenishment solution requires very frequent updates of data, as well as relatively high data accuracy levels in order to be the most effective. An axiom in inventory management applies: the lower the inventory level, the better the accuracy and management of inventory must be.
- Software helps standardize processes. It formalizes the processes required for data collection and processing. However, some complexities exist with software; the IT department must cooperate fully. For example, if IT insists on using a self-made solution rather than a ready-made system, the implementation might be delayed by several months and sometimes even years. A proper financial justification focusing on the increase in T and reduction in I investment based on the rapid implementation of a turnkey system is hard to deny.

Let me elaborate on this last point. In order to be able to make the project a success, three groups in the company must cooperate with each other. They are:

- The owner
- The end users
- The IT staff

Each has their own goals, needs, and wants. Each must buy-in to the solution for it to be implemented and managed effectively.

- The owner is defined as the top decision maker in the organization. Since the TOC pull distribution/replenishment solution requires changes that typically require high-level authorization, it is best to convince the top decision maker to embark on the TOC solution. Customarily, the owner's goal is to get the best financial results possible for the organization, as his personal compensation (if he is not the owner himself) will usually be measured this way as well. Therefore, to convince the owner to embark on a TOC project is relatively straightforward as TOC directly targets financial results through the T channels. However, the owner plays another, more important role in the TOC project; by demonstrating his personal and active involvement in the project and radiating the project's top priority to everyone in the organization, the owner can make the project very successful. Without his championing the project, the project might suffer from lack of others' attention, dragging the implementation out for months, and eventually leading to poor results.

- End users implement and manage using the new processes demanded by the TOC pull distribution/replenishment solution. The end user includes “anybody who performs an action with or according to the software.” The end users, while less effective as individuals, are very influential as a group. Therefore, the buy-in process here is no less important. Proper training must be conducted, explaining the reasons for the change of processes.³⁵ Users can buy-in to the concepts but not perform the required processes to make it happen. The end user goal is convenience, and it is important to educate the end users and thoroughly explain why after switching to the new methodology the end user’s life would become much simpler, not more complex, especially when he will be able to have much better control in meeting his job responsibilities and much higher visibility in the organization. It is also very important to create a win for the end users by reaching the goal of the implementation—the best way is to tie the success of the project to their own income by bonuses, stock options, or otherwise. This will ensure they will be totally committed to the changes they need to endure and will willingly embrace them.
- The IT staff function is central to any big initiative, especially the TOC distribution solution. IT has a major impact on the start of the project. They must install the software and make sure it is running correctly. They must import or link the software to inputs and outputs, and then later perform maintenance and updates to the software. IT can really affect the parameters and are usually the internal experts in the company for setting the correct parameters in the software. The IT function is generally filled by very capable, educated people who are very analytical and intelligent. Therefore, if IT is

correctly bought in, they will become very powerful allies to the implementation and will help achieve the full project benefits. The main goal of the IT function is typically self-value. In order to achieve the buy-in, it is important to communicate to IT that their influence will grow after implementing TOC (since the organization will heavily depend on them for making parameter decisions). Therefore, IT is key to the success of the implementation.

Actual Results of the TOC Distribution/Replenishment Solution

Based on the combined experience of TOC consultants and software companies in implementing the TOC distribution/replenishment solution,³⁶ it is safe to say that the results are remarkable. Using the approach listed here (especially to set initial buffer sizes), significant results were achieved in three months. The average results of implementing the TOC solution are a 40 percent increase in sales, coupled with a 50 percent reduction in inventory investment. Inventory turns improved by a factor of 2.8. Think about the impact of this solution on the ROI in inventory.

These impressive results demonstrate that the TOC distribution/replenishment solution works. Planning the implementation carefully, selecting the right consultants, selecting the right software, and creating buy-in are requirements; successful management causes a huge competitive edge, increased control over inventory and sales, and therefore higher profitability.

Summary

It is clear that traditional supply chains do not function effectively. Most organizations have given up on the possibility of having 95 percent or higher inventory availability. If organizations do reach a state of 95 percent or higher availability, they do it with huge inventories and the associated cost of keeping excess inventory (and where inventory is missing, the high expediting costs). On the other hand, stockouts hurt their sales as well. The dilemma is whether to stock little inventory (and suffer stockouts and lost sales) or to stock a lot of inventory (and suffer the high inventory investment and associated inventory costs). Recall that we must have the right item (what) at the specific location (where) at the right time (when) to be successful. It is clear that if an effective and simple solution exists to answer these questions without having large inventories, then organizations would willingly embrace it.

The TOC distribution/replenishment solution is quite new in comparison to the reorder point/economic order system invented by Harris (1915) and the min-max inventory system (the basic models used in many distribution requirements planning systems) invented shortly thereafter. In comparison to these inventory systems, the TOC system is the new kid on the block. A major fundamental of the TOC system is the use of the PWH as the hub in the distribution network. PWHs³⁷ have existed in the past but were not considered the major distribution point or the buffer protecting the whole network. PWHs held little inventory. This centralization of inventory at the PWH concept is now making a comeback (this time called “logistical centers”) but the understanding that the system would function, financially and operatively, much better by pull rather than by push is relatively new. The TOC solution uses the PWH/CWH as the hub and pulls inventories through the chain to the consumption point. This pull approach is new. The TOC concepts of stock buffer size, BM, the focus on T,

and the DBM mechanism are new, unique, and very effective. TOC offers remarkable results achieved in a short time period, often in a manner of “too good to be true.” Implementation of the TOC distribution/replenishment solution is difficult (it is a paradigm shift) but, by following a few simple guidelines, the obstacles can be minimized.

References

Blackstone Jr., J. H. 2008. *APICS Dictionary*. 12 ed. Alexandria, VA: APICS.

Cox III, J. F. and Walker II, E. D. 2006. “The poker chip game: A multi-product, multi-location, multi-echelon, stochastic supply chain network useful for teaching the impacts of pull versus push inventory policies on link and chain performance,” *INFORMS Transactions on Education, Special Issue on Supply Chain Management Education* 6(3):3-19.

Harris, F. E. 1916. “What quantity to make at once,” *The Library of Factory Management*. Vol. V. Operations and Costs. Chicago: A. W. Shaw Company, pp. 47-52.

Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>.

Recommended Reading

Goldratt, E. M. 2009. *The Choice*. Great Barrington, MA: North River Press.

Knowledge Base at: www.inherentsimplicity.com.

Schragenheim, E., Dettmer, H. W., and Patterson, J. W.
2009. *Supply Chain Management at Warp Speed*. Boca
Raton, FL: CRC Press.

About the Author

Amir Schragenheim, since 2004, has been the President of Inherent Simplicity Ltd., a software firm specializing in TOC software for Production and Distribution environments. Inherent Simplicity is currently the only software supplier in the areas of production and distribution to Goldratt Consulting, Eli Goldratt's consulting firm, in their Viable Vision strategic projects.

Mr. Schragenheim holds an MBA from Tel Aviv University magna cum laude, and majored in Marketing and Strategy. He is a TOCICO certified expert in Supply Chain Logistics, Project Management, Finance and Measures and Holistic Business Strategy. Mr. Schragenheim is a regular speaker at both the TOCICO International and Regional Conferences.

Mr. Schragenheim started his professional career with TOC in 1998 when with Eli Schragenheim, he developed computer simulations of production and project management to demonstrate the power of TOC.

CHAPTER 12

Integrated Supply Chain Beyond MRP—How Actively Synchronized Replenishment (ASR) Will Meet the Current Materials Synchronization Challenge

Chad Smith and Carol Ptak

Introduction

The effectiveness of any system has to be judged by the results that it achieves. In today's environment, companies and supply chains that struggle with effective materials planning consistently see at least one or a combination of three main business results:

- **Unacceptable inventory performance.** This is identified as having too much of the wrong material, too little of the right material, high obsolescence, or low inventory turns. Companies frequently can identify many of these problems at the same time.
- **Unacceptable service level performance.** Customers continue to put pressure on the company, which quickly exposes poor on-time delivery, low fill rates, and poor customer satisfaction. In addition, customers consistently attempt to drive prices down.

- **High expedite-related expenses and waste.** In an attempt to fix the previous two unacceptable business results, managers will commit to payment premiums and additional freight charges or increase overtime in order to fulfill promises. When the promises are still not fulfilled, then the company is exposed to financial penalties.

The purpose of this chapter is to present an alternative demand-driven approach for planning and controlling material flow and contrast it to the poor business results embedded in most traditional *material requirements planning* (MRP) systems. This includes a discussion of the core problems causing these results. The concepts and procedures underlying this new planning and control system are based on several *Theory of Constraints* (TOC) concepts including strategic buffering, replenishment, and *Buffer Management* (BM).

Copyright © 2010 by Chad Smith and Carol Ptak.

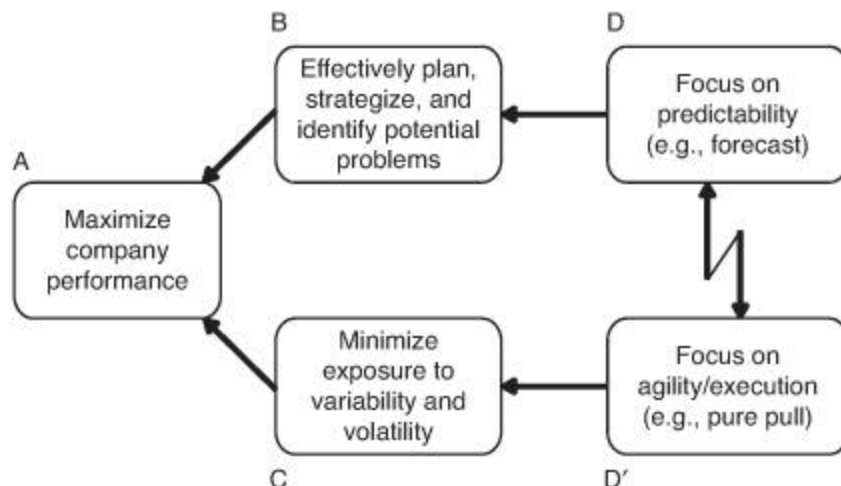


FIGURE 12-1 The current situation for many complex manufacturing environments.

Actively Synchronized Replenishment (ASR) is not dependent on a *Drum-Buffer-Rope* (DBR) environment, but many DBR implementations will be dependent on ASR. The same is true for Lean environments. ASR is not dependent on Lean, but many Lean implementations will benefit from the implementation of ASR. Both DBR and Lean are pull systems and are inherently in conflict with the standard material planning systems, which push materials. This demand-driven materials and inventory approach is in many ways agnostic to a company's desired capacity scheduling approach. In other words, no matter what kind of capacity scheduling approach a company chooses to use, a methodological compromise is not required to ensure material availability. This chapter provides the description of a proven approach that successfully creates pull-based materials flow and synchronization in complex environments where traditional MRP was historically a necessity but performed its functions poorly.

The conflict cloud in [Fig. 12-1](#) clearly describes the current situation for many complex manufacturing environments. On one hand, there is a necessity to effectively plan in advance of real customer orders to order long lead time materials, incorporate sales and marketing data and plans, plan capital and staffing levels, and develop contingency plans for potential problems. This has driven the management team to focus on systems and approaches that emphasize predictability. Some companies have developed a very sophisticated sales and operations planning process in order to minimize the potential for problems within the planning horizon. On the other hand, there are three well-known rules of forecasting.

1. Forecasts are always in error.

2. The more detailed the forecast is, the more error will be realized.
3. The further into the future the forecast goes, the more error will be realized.

These three rules of forecasting represent how the focus on predictability exposes companies to the risk associated with variability and volatility. The necessary inventory and resource costs to compensate for forecast error are too expensive in this hypercompetitive time. This has driven managers to focus on reducing planning lead times and implementing pull-based strategies like Lean and DBR to improve overall company agility. It is well known that when a company can react quickly, then there is less exposure to market volatility and variability. In order to resolve this conflict effectively, a solution must be deployed that allows companies to effectively plan and strategize without the inherent risks that come along with conventional approaches.

The organization of this chapter consists of this introduction which briefly describes the realities of manufacturing complex long lead-time products in a constantly changing environment. Next, we surface problems (undesirable effects) and then identify the underlying cause (core problem) of using push systems to manage both production and inventories in this environment. Finally, the direction and exact solution to the core problem is described. To demonstrate how significant this approach has been, some case studies of implementation successes are presented.

Identifying the Real Problem—Rethinking the Scope of Supply Chain Management

In the last 20 years, there has been much attention and emphasis on developing supply chain solutions from both a methodological and a technological perspective. In truth, most of what has been developed has been a revolution for the distribution and logistics between consumers and suppliers. Distribution and logistics are no longer the constraint worldwide. Now it is well known across the supply chain what is sold and when it has moved. A logistics company can provide real-time updates as parts move around the world. Ultimately, however, at the heart of any supply chain is manufacturing and, in most supply chains, it is several different manufacturing sites and processes that must be effectively coordinated and synchronized to bring a finished item into the distribution pipeline. Now, the question is how to increase the coordination and synchronization? An AMR Report concluded that:

Today's companies have a dilemma. They need to reduce costs in the face of product complexity, shorter product lifecycles, and increased regulatory compliance. While companies apply a broad range of supply chain strategies to address these challenges, the buck ultimately stops with manufacturing. This is forcing a fundamental redefinition of the role that manufacturing needs to play in today's supply networks, underscoring the need for demand-driven manufacturing and agility. (Masson et al., 2007, 1)

In truth, *Supply Chain Management* (SCM) solutions do not deal with the manufacturing implications and coordination (materials and capacity) of the items that they are demanding and supplying. While there is a wide array of different (and effective) methodologies and technologies to schedule manufacturing capacity, there is one universal materials system and approach throughout the world to manage materials known as MRP. To be consistent with

current global understanding we will use the following definition from the *APICS Dictionary* (Blackstone, 2008, 81):

Material requirements planning (MRP)—A set of techniques that uses bill of material data, inventory data, and the master production schedule to calculate requirements for materials. It makes recommendations to release replenishment orders for material. Further, because it is time-phased, it makes recommendations to reschedule open orders when due dates and need dates are not in phase. Time-phased MRP begins with the items listed on the MPS and determines (1) the quantity of all components and materials required to fabricate those items and (2) the date that the components and material are required. Time-phased MRP is accomplished by exploding the bill of material, adjusting for inventory quantities on hand or on order, and offsetting the net requirements by the appropriate lead times. (© APICS 2008, used by permission, all rights reserved.)

Let us be very clear—MRP is not going away (and it shouldn't). Since the detailed description of its specifications by Orlicky (1975) in his classic book, *Material Requirements Planning*, MRP has provided the foundation for the design and material planning within most manufacturing environments. An Aberdeen Group Study (2006, 17, Table 3) showed that 79 percent of companies that bought *Enterprise Resources Planning* (ERP) systems also bought and implemented the MRP module.

Even after over 50 years of using MRP and other technologies to plan and coordinate material, how is it that companies and supply chains can struggle so mightily with materials synchronization and the business effects identified at the start of this chapter? After careful examination of many companies and the supply chains in which they

participate, there appear to be two main reasons why those effects happen in today's manufacturing enterprises:

1. **MRP was not designed to deal with today's challenges.** The sheer size of ERP systems today hides the reality that for most mid-range and large manufacturers, MRP remains a critical module in their ERP system, and the changing global manufacturing environment has exposed critical shortcomings in most MRP implementations and tools. Variability and volatility are on a dramatic rise and the implementations of pull-based philosophies like Lean and TOC are proliferating. These conditions and approaches are putting extreme pressure on MRP systems and even creating conflicting modes of operation (push versus pull). We need to be reminded that MRP was designed in the 1950s, commercially coded in the 1970s, and really has not changed since. The reality is that it was never designed with today's factors and pull-based concepts in mind.
2. **Users are forced to make incomplete and unsatisfactory compromises.** Most companies are not blind to the shortcomings. Materials and Production Control personnel often find themselves in a dilemma regarding their MRP system. There are powerful aspects of MRP that are still relevant and necessary. MRP is possibly even more relevant than ever as we have more complex planning scenarios than ever. At the same time, there are disastrous consequences to ignoring MRP's shortcomings in today's environment. Given this conflict, Materials and Production Control personnel are forced to find various, often unsatisfactory and incomplete, ways around this conflict.

A Brief History of MRP

The invention of MRP in the 1950s was nothing short of a revolution for manufacturing. For the first time, companies could plan for needed materials based on an overall master schedule exploded through a *bill of materials* (BOM). The manual single- and double-order point systems were no match for the proliferation of products coming to market after World War II. The world was in the age of marketing! We found that we could no longer live without things that did not exist 10 years earlier. Class “A” MRP implementations yielded significantly reduced inventory and improved on-time deliveries. APICS—the American Production and Inventory Control Society—was founded in 1957 in Ohio to disseminate the education necessary to effectively use the tools that were quickly being developed. In 1976, the APICS CPIM certification was introduced and quickly became a standard worldwide of the mastery of the production and inventory control techniques of the day including inventory management, MRP, production activity control, and master planning. Driven by this available APICS education through the 1970s and with the APICS MRP Crusade, MRP quickly became the number one tool that inventory-related management personnel relied upon to ensure that material was available to meet manufacturing and market requirements.

Even in these simpler, more predictable times, MRP was successful as measured by significant bottom-line results including dramatic inventory reduction in only a small percentage of companies that implemented the tool. The early adopters showed significant results, but as MRP came into more widespread use, the same results were not achieved. This significant failure rate of MRP was a major point of discussion in the APICS meetings at the time. One big reason was that MRP was intended to do only that—plan

material. APICS professionals at the time knew that capacity was a critical consideration. However, the computer power at the time was limited and even if the capacity algorithms were available, it was just not possible to calculate both at the same time. Remember that the first MRP systems were written in only 8K of memory! However, computers quickly became more powerful and closed-loop MRP was developed to answer the problems of the day. The *APICS Dictionary* (Blackstone, 2008, 21) defines *Closed-loop MRP* as:

A system built around material requirements planning that includes the additional planning processes of production planning (sales and operations planning), master production scheduling, and capacity requirements planning. Once this planning phase is complete and the plans have been accepted as realistic and attainable, the execution processes come into play. These processes include the manufacturing control processes of input-output (capacity) measurement, detailed scheduling and dispatching, as well as anticipated delay reports from both the plant and suppliers, supplier scheduling, and so on. The term closed loop implies not only that each of these processes is included in the overall system, but also that feedback is provided by the execution processes so that the planning can be kept valid at all times. (© APICS 2008, used by permission, all rights reserved.)

Closed-loop MRP was the next evolution and allowed the planning of both material and capacity. Still, the development and implementation of an MRP system was far from a guarantee of success. The tool was far more sophisticated and the available APICS education provided people who understood how the tools worked, but still the implementation was not a guarantee of success.

Technology became more powerful and the client-server age was upon us. In the 1980s, MRP II (manufacturing

resources planning) was developed to provide further integration to the core business system by incorporating the financial analysis and accounting functions. *MRP II* in the *APICS Dictionary* (Blackstone, 2008, 78) is defined as:

A method for the effective planning of all resources of a manufacturing company. Ideally, it addresses operational planning in units, financial planning in dollars, and has a simulation capability to answer what-if questions. It is made up of a variety of processes, each linked together: business planning, production planning (sales and operations planning), master production scheduling, material requirements planning, capacity requirements planning, and the execution support systems for capacity and material. Output from these systems is integrated with financial reports such as the business plan, purchase commitment report, shipping budget, and inventory projections in dollars. Manufacturing resource planning is a direct outgrowth and extension of closed-loop MRP. (© APICS 2008, used by permission, all rights reserved.)

MRP II systems became more commercially available. No longer was it necessary for companies to develop these systems. Software companies catering to the needs of different industries and platforms provided a wide variety of software products off the shelf. At the same time, the APICS education and certification program provided industry with professionals capable of utilizing these systems. Still, the systems that were so advanced at the time were no guarantee of bottom line success. In the 1990s, as technology began to move to Internet architecture, ERP was the next evolution and brought all the resources of an enterprise under the control of a centralized integrated system. In the *APICS Dictionary*, *ERP* (Blackstone, 2008, 45) is defined as:

Framework for organizing, defining, and standardizing the business processes necessary to effectively plan and control an organization so the organization can use its internal knowledge to seek external advantage. (© APICS 2008, used by permission, all rights reserved.)

Companies continued to invest in technology pursuing the holy grail of integrated planning and yet significant bottom line results were not achieved. In the mid 1990s, advanced planning and scheduling (APS) systems¹ leveraged the visibility of the company's resources in ERP and promised to keep all scarce resources busy all the time. The *APICS Dictionary* (Blackstone, 2008, 4) defines an APS as:

Techniques that deal with analysis and planning of logistics and manufacturing during short, intermediate, and long-term time periods. APS describes any computer program that uses advanced mathematical algorithms or logic to perform optimization or simulation on finite capacity scheduling, sourcing, capital planning, resource planning, forecasting, demand management, and others. These techniques simultaneously consider a range of constraints and business rules to provide real-time planning and scheduling, decision support, available-to-promise, and capable-to-promise capabilities. APS often generates and evaluates multiple scenarios. Management then selects one scenario to use as the "official plan." The five main components of APS systems are (1) demand planning, (2) production planning, (3) production scheduling, (4) distribution planning, and (5) transportation planning. (© APICS 2008, used by permission, all rights reserved.)

Once again, the implementation of these complex systems was rarely a significant bottom line success. This is not to say that the software did not implement or did not run. The reality was that the improved bottom line results

promised in the business case were the exception rather than the rule.

Throughout this entire evolution, the MRP calculation kernel stayed the same. Fundamentally, MRP is a very big calculator utilizing the data about what you need and what you have to calculate about what you need to go get and when. At its very core, even the most sophisticated ERP system of the day is inherently a push system based on a forecast or plan and the assumption that all the input data are accurate. In the most stable of environments, this assumption may be somewhat possible, but how does the 21st century global economic environment fit with this approach?

Can MRP Meet Today's Challenge?

The world that existed when MRP was developed no longer exists. We are now in a world where global capacity far exceeds global demand. Customers can purchase what they want, when they want it, at a price they want to pay due to the lack of transactional friction available now through the Internet. Since they now have this freedom to go anywhere to purchase anything with a few clicks of a mouse, customers are increasingly fickle. The push strategy of produce and promote of the post-WWII era just does not work anymore.

While some manufacturers turn to various technologies and process improvement approaches to reduce variability in individual processes on the shop floor, the reality is that variability and volatility are rising dramatically when you examine the bigger picture. No longer can a company compete simply by looking internally. Now a company must consider the entire enterprise as well as the supply chain within which it operates. Today's manufacturing operations

are far more susceptible to disruptions throughout their internal operations and external supply chain due to:

- Global sourcing and demand
- Shortened product life cycles
- Shortened customer tolerance time
- New materials
- More product complexity and customization
- Demands for leaner inventories
- Inaccurate forecasts
- Material shortages
- Complex synchronization issues
- More product variety
- Long lead time parts/components
- More offshore suppliers

The bottom line is that these factors combine to create an environment where planning scenarios that are more complex exist and those scenarios often come with higher stakes attached.

In [Table 12-1](#), we outline the organizational effects of typical MRP implementation attributes.

Typical MRP Attributes		MRP Effects to the Organization
Planning Attributes	MRP uses a forecast or master production schedule as an input to calculate parent and component level part net requirements.	Part planning becomes based on a "push" created by these forecasted demand requirements. Forecast accuracy at the individual SKU and part levels is highly inaccurate. Build plans and PO's that are calculated from this forecast often are misaligned with actual market demand. This leads to excessive expediting, overtime, premium freight, increased inventory of the wrong items, and missed shipments.
	MRP pegs down the entire BOM to the lowest component part level whenever available stock is less than exploded demand.	Creates a complicated materials and scheduling profile that can totally change with one small change at a parent item. When capacity is scheduled infinitely, there are massive priority conflicts and material diversions. When capacity is scheduled finitely across all resources, there is massive schedule instability due to cascading slides from material shortages.
	MRP allows the release of work orders to the shop floor without consideration of component parts availability.	MOs are released to the floor but cannot be started due to shortages. This leads to increased WIP, constantly changing priorities and schedules, delays, expediting, and overtime.
Stock Management Attributes	Lead time for parent part is the manufacturing lead time only for the parent, regardless of the cumulative lead time for parent and lower level component parts.	MOs are often released with dates that are impossible to achieve or without all component parts available.
	Fixed reorder quantity, order points, and safety stock that do not adjust to actual market demand or seasonality.	Additional exposure to forecast inaccuracies resulting in increased expediting.
	Only parts hitting minimum or reorder point are flagged for reorder.	Aggregate inventory visibility is limited, frequently putting the company in a constant expedite mode. Additionally, there is no way to judge relative priority between stock orders.
	OPTION 1: Past due requirements and orders to replenish safety stock are often treated as "Due Now."	Every safety stock order looks the same, which means there is no real priority. To determine real priorities requires massive attention, analysis, and priority changes.
	OPTION 2: Priority of orders is managed by due date.	Due dates will not reflect actual priorities. To determine real priorities requires massive attention, analysis, and priority changes.
	Limited future demand qualification. Limited early warning indicators of potential stockouts or demand spikes.	Planners either have to bring in all future demand, which inflates inventories and wastes capacity and materials, or have to bring in no future demand, which makes the environment extremely vulnerable to spikes or must pour through large amounts of data in order to qualify spikes for each part.

Table 12.1 The Organizational Effects of Typical MRP Attributes

As defined in the *APICS Dictionary* and taught in APICS education, these basic MRP attributes and functions are well understood. The limitations and implementation issues have been the subject of many APICS dinner meetings and conference presentations over the lifetime of the technology. One only needs to examine the APICS international conference proceedings from the past three decades and you will discover a variety of proposed solutions and workarounds. The early pioneers like Ollie Wight, George Plossl, Dave Garwood, and Walt Goddard provided many ideas that were built upon as practitioners continued to struggle with these issues. These suggested policies, procedures, and workarounds, however, can contain functionality that has nothing to do with MRP. Sometimes this additional functionality simply moves the pain points to another part of the organization. Many times, the additional functionality does not overcome limitations

that are more fundamental and design issues that tend to go unaddressed.

Conventional MRP implementations just do not fit the new pull-based manufacturing and materials solutions required to be fast, lean, and flexible in today's hypercompetitive environment. Users are frustrated because they cannot complete their work inside the system. To get the job done, they extract data to Excel[®] or Access[®]. Even worse, they use manual sticky notes and manual scheduling white boards. Gone is the desired integration driving the investment in the formal system. In the effort to get the job done at any level, the IT landscape is more complicated and the costs to support it constantly increase.

The MRP Conflict Today

Does your company work within its formal planning system or does your company work around this system? Does it try to do both at the same time? Are spreadsheets, sticky notes, and manual tracking systems still alive and well in your operations even though you have implemented an MRP or ERP system in the last 10 years?

When it comes to truly effective materials management, most Purchasing, Manufacturing, and Production Control personnel frequently feel like their hands are bound and tied. MRP's power has always been its ability to manage BOM connections in order to generate total net material requirements (demand orders that turn into manufacturing orders or purchase orders). The more complex and integrated the product structures, manufacturing facilities, and supply chains are, the more necessary MRP is for netting and getting ahead of critical and long lead time parts. Most Purchasing, Manufacturing, and Production Control personnel realize this and are forced into a set of unsatisfactory compromises that just don't work. The next

section discusses the compromises that arise from this MRP conflict.

The MRP Compromises

In most cases, there are five types of compromises that frequently occur (either separately or in combination).

- 1. Manual Work Around Proliferation**—As has been discussed already, companies frequently try to work around their MRP system by relying on stand-alone, disconnected, and highly customized data manipulation tools like Excel spreadsheets and Access programs. These tools have serious limitations and their proliferation makes the IT landscape more complicated and maintenance more intensive. Their use ultimately defeats the purpose behind the major investment in an integrated ERP package.
- 2. Flatten the BOM**—Sometimes companies try to simplify the synchronization issue by flattening the BOM. Flattening the BOM removes levels that were originally identified to define the product and the process. The key to better synchronization is not to ignore dependencies within the product structure and across product structures. Better synchronization is possible when you know on which dependencies to focus. When the BOM is flattened, it is imperative that only those BOMs are flattened that cannot provide a leverage point. Flattening BOMs across the board can eliminate key leverage points that can provide a great deal of value. These dependencies provide an excellent way to stop variability from gaining momentum and disrupting the entire supply chain like a tsunami wave. The key to better synchronization is

to understand those dependencies and control them. By flattening the BOM, companies can actually lose visibility at both the planning and the execution levels. In some cases, companies can benefit themselves by inserting an additional level in the BOM!

3. **Make-to-Order Everything**—Still other companies choose to place all of their cash in raw material and purchased components and embrace a completely *make-to-order* (MTO) strategy. In most environments, this comes with a significant price. A company either has to carry additional capacity to meet service level requirements or risk service level satisfaction with extended lead times. In some highly seasonal or short customer tolerance environments, this is simply impossible. The company just cannot supply the product in sufficient time with sufficient volume.
4. **More Efficient Forecasting**—Other companies implement advanced forecasting algorithms or hire more planners in hopes of guessing better. Recall that the assumption under MRP is that there is a plan or forecast that is the demand in the system to drive the MRP calculation. Even with dramatic improvements in forecasting accuracy, the results do not translate to the bottom line. Experience has shown that at best these solutions result in a 20 to 40 percent improvement in demand signal accuracy—still leaving significant room for error. Even if a company succeeds in increasing signal accuracy, it does not necessarily translate well to overall effectiveness in terms of availability and fill rates. Remember, the increase of variability and volatility (especially on the supply side) can easily offset any appreciable gain in signal accuracy. Also, remember that many manufacturers can have multiple assembly and subassembly

operations that are integral parts of their overall flow. In any type of assembly operation, it takes the lack of only one part to block a complete shipment. The more assemblies there are, the more complex the synchronization and execution challenge is. Finally, even the biggest supporters of forecasting cannot argue the fact that forecasting in any form is still a push-based tactic. Yes, it can be a more educated push but it is still a push nonetheless. For companies implementing pull-based manufacturing systems (e.g., Lean or DBR), this sets up conflicting modes of operation that will simply not perform well in volatile and complex environments.

- 5. Manual Reorder Point Systems**—With the implementation of kanbans, supermarkets and three-bin systems manufacturing have come full circle. Unable to overcome the shortcomings associated with MRP, some companies have abandoned it completely. It is essentially throwing the proverbial baby out with the bath water. In many environments, it is devastating. These systems tend to be manually intensive and very difficult to make responsive to changes in the environment. There is almost no ability to see either the truly available stock or the total net requirements picture (all demand allocations in relation to all open supply orders). Real data is masked in traditional systems by requirements coming from forecasts or other false demand signals. In fact, by definition, each parent-child relationship in the BOM is managed independent of any other connection. MRP consolidates the total requirements for each child part and only rarely can even an experienced planner understand why that quantity is being ordered. In environments with high variety and options, it often requires massive amounts of

inventory on the floor in order to be able to provide components and parts when necessary. A 2007 AMR Report (Masson et al., 2007, 6) came to two important conclusions. First that, “Kanban cards and heijunka boards become unmanageable when there are hundreds or thousands of products and components.” Second, and most interesting, is that in large global manufacturers with many manufacturing sites and lines, “The pragmatist needs software to support lean manufacturing.” Remember that simply knowing the stock on hand cannot provide the information to know what to order unless the on-hand position plus the open supply orders minus the demand allocations is considered (this is called an available stock equation). This is just not possible with manual reorder point systems like kanbans.

Actively Synchronized Replenishment—the Way Out of MRP Compromises

For those who are familiar with constraints management and its thinking processes, the dilemma that manufacturing companies find themselves in can be seen in the conflict cloud in [Fig. 12-2](#).

There are essentially two critical needs (B: Produce to demand and C: Visibility to total requirements) coming into contention behind the compromises (made at D and D', the pull or push choices). From a manufacturing perspective, we must have a realistic way to respond and produce to demand. This way must include both capacity *and* materials. MRP tools simply do not create the correct “demand signals,” nor do they facilitate materials availability within increasingly shorter horizons that are inherently more variable and volatile. Additionally, many pull-based manufacturing implementations (e.g., Lean and

DBR) are effectively blocked by this lack of material synchronization. In most cases, due to the shortcomings listed previously, this leads many manufacturing personnel within companies to think that they should ignore MRP. In fact, a frequent milestone for a Lean implementation is that the computer planning system has been eliminated!

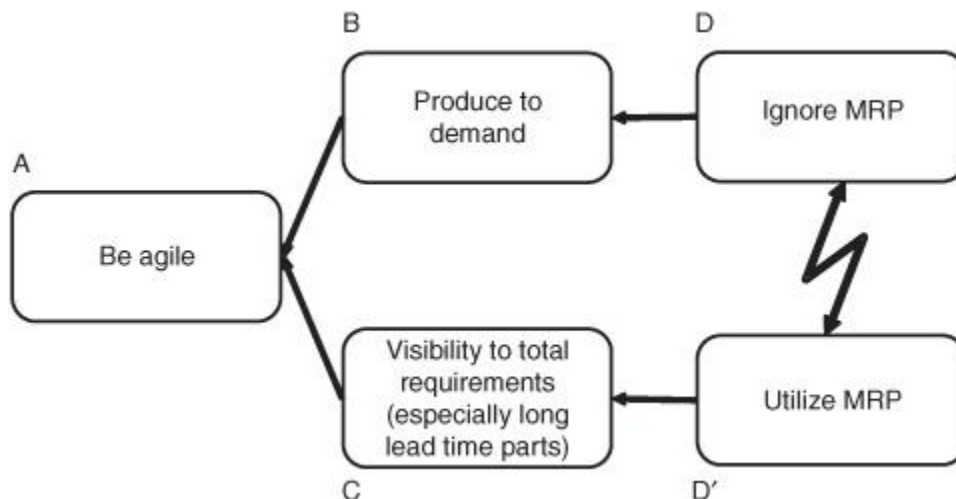


FIGURE 12-2 The conflict in utilizing MRP.

On the other hand, from a Planning and Purchasing perspective we must have a way to effectively see, plan, synchronize, and manage the availability of *all* materials, components, and end items, especially critical and long lead time manufactured and purchased parts. With increasingly complex planning scenarios, it leads Planning personnel to insist on utilizing MRP.

The more complex the manufacturing environment, the more acute this conflict tends to be. The inability to effectively reconcile the dilemma in those environments leads to the ineffective MRP compromises listed previously and can also essentially relegate TOC, Lean, and Six Sigma implementations to lip service. The requirements must be achieved without the conventional inaccuracy, inconsistency, and massive additional efforts and waste associated with the current set of compromises. MRP, as

previously noted, has some very valuable core attributes in today's more complex planning and supply scenarios (BOM visibility, netting capability, and maintenance of sales order/work order connection between demand allocations and open supply). The key is to keep those attributes but eliminate MRP's critical shortcomings (listed previously) and use the pull-based replenishment tactics and visibility behind TOC and Lean concepts all in one system in a dynamic and highly visible format.

ASR builds upon the traditional replenishment approach of replacing what was taken or used to create a dynamic and effective pull-based solution to answer the challenges of today's manufacturing landscape. Through new approaches in inventory and product structure analysis, new pull-based demand planning rules, and integrated execution tactics, ASR is designed to directly tie material availability and supply to *actual* consumption throughout the BOMs, thus removing the "islands of MRP" obstacles that most supply chains face. Additionally, this approach is a prerequisite to effectively utilize pull-based scheduling and execution methods like Lean and DBR in more complex manufacturing environments. Additionally, ASR has a unique way to incorporate required elements of strategic planning with little or no exposure to the variability and volatility that gets companies in trouble with traditional forecasting techniques.

ASR has five main components.

1. Strategic Inventory Positioning
2. Dynamic Buffer Level Profiling and Maintenance
3. Dynamic Buffers
4. Pull-Based Demand Generation
5. Highly Visible and Collaborative Execution

They are discussed in the next sections.

1. Strategic Inventory Positioning

The first question of effective inventory management is not, “How much inventory should we have?” The most fundamental question to ask in today’s manufacturing environments is, “given our system and environment, where should we place inventory to have the best protection?” Think of inventory like a break wall to protect boats in a marina from the roughness of incoming waves. Out on the open ocean, the break walls have to be 50 to 100 feet tall, but in a small lake the break walls are only a couple of feet tall. In a glassy smooth pond, no break wall is necessary.

In the same way, inventory is the break wall against the variability experienced from either supply (externally and internally) or demand unreliability. Remember that a company has to think holistically across not only the enterprise but also across the supply chain. Putting inventory everywhere is an enormous waste of company resources. Eliminating inventory everywhere puts the company and supply chain at significant risk. Strategically positioning inventory ensures the company’s ability to absorb expected variability without having to disrupt every part of the plant and the supply chain. Important factors to consider carefully in determining where to place inventory buffers include:

- **Customer Tolerance Time**—The time the typical customer is willing to wait or the potential for increased sales for lead time reductions.
- **Variable Rate of Demand**—The potential for swings and spikes in demand that could overwhelm resources (capacity, material, cash, credit, etc.).

- **Variable Rate of Supply**—The potential for and severity of disruptions in particular sources of supply or specific suppliers.
- **Inventory Flexibility and Product Structure**—The places in the “aggregate BOM” structure that leave a company with the most available options (primarily key purchased materials and subassemblies/components). *The aggregate BOM structure can be defined as the holistic BOM across the company with all identified product interrelationships.* The more shared components and materials there are, as well as the deeper and more complex the aggregate BOM is, the more important this factor is. Through a process known as *BOM decoupling*, variability is absorbed, cumulative lead times are compressed and reduced, and planning is simplified by the insertion of ASR buffers at these strategic points in the BOM. What is important to note is that decoupling should not occur at every connection in the BOM, only at the connections that really make the biggest impact (more on this later). By combining the aggregate BOM concept with the BOM decoupling concept, key child components that compress the lead times of the most parents can be identified. In addition, currently stocked positions that do not truly compress lead times for parents can be identified and eliminated. We will explore this in depth in the section on ASR planning.
- **The Protection of Key Operational Areas**—It is particularly important to protect critical operational areas from the bullwhip effect. The *bullwhip effect* is the cascading disruptions through a dependent sequence of events. This undesirable effect of MRP and push distribution systems is well known. The *APICS Dictionary* (Blackstone, 2008, 15) defines *bullwhip effect* as “(a)n extreme change in the supply position upstream in a supply chain generated by a small change in demand

downstream in the supply chain. Inventory can quickly move from being backordered to being excess. This is caused by the serial nature of communicating orders up the chain with the inherent transportation delays of moving product down the chain. The bullwhip effect can be eliminated by synchronizing the supply chain.” (© APICS 2008, used by permission, all rights reserved.)

Within manufacturing, it can be eliminated by synchronizing the pull across the production processes. MRP does not do this.

The longer and more complex the routing structure and dependent chain of events (including inter-plant transfers), the more important it is to protect key operations. These types of operations include areas that have limited capacity or where quality can be compromised by disruptions. In some cases, the creation of new part numbers and an insertion of an additional level in the BOM (as opposed to deleting layers) are necessary in order to decouple long and complex routings or sequences.

These factors are applied across the entire BOM and supply chain to determine positions for purchased, manufactured, and subcomponents and finished items (including service parts). Purchased parts chosen for strategic replenishment tend to be critical or strategic parts and long lead time items. Typically, this will be less than 20 percent of purchased parts. Manufactured parts chosen for strategic replenishment are often critical or strategic manufactured and service parts, at least some finished items, and critical subassemblies.

The position of these buffers is accomplished through a combination of “thoughtware” and software. The “thoughtware” is the application of most of the above factors in consideration of the business objectives and operating rules by the people that have experience and intuition in the environment. In complex environments, software is often required to do the heavy computational lifting in order to analyze product structure, cumulative lead times, and shared components across the aggregate BOM. Finally, the importance of this step should not be underestimated. Without the right strategic positioning, no inventory system can live up to its potential.

2. Dynamic Buffer Level Profiling and Maintenance

Once the strategic inventory positions are determined, the actual levels of those buffers have to be initially set. Based on several factors, different materials and parts behave differently (but many also behave nearly the same). ASR groups parts and materials chosen for strategic replenishment and that behave similarly into “buffer profiles.” Buffer profiles take into account important factors including lead time (relative to the environment), variability (demand or supply), and whether the part is made or bought. For instance, you could have a group of purchased parts that are long lead time and high variability (subject to frequent disruptions in supply) and you could have a group of manufactured parts that are short lead time and high variability (subject to frequent spikes). These buffer profiles produce a unique buffer picture for each part as their respective individual part traits are applied to the group traits. A list of both group and individual part traits that can apply to create that unique picture for every part is given in [Table 12-2](#).

Group Trait Examples	Individual Part Trait Examples
<ul style="list-style-type: none"> • Lead Time (Long, Medium, Short) • Make or Buy • Supply Variability (High, Med, Low) • Demand Variability (High, Med, Low) 	<ul style="list-style-type: none"> • Average Daily Usage (ADU) • Fixed Lead Time • ASR Lead Time • Minimum-Order Quantity • Maximum-Order Quantity • Order Multiple • Seasonality

TABLE 12-2 Part Trait Examples

This unique buffer picture is not just what the top level quantity should be. In [Fig. 12-4](#), we see that ASR stratifies the total buffer level into different “zones.” ASR uses a five-colored zone stratification approach. Light blue (LB; some authors refer to this as the white zone) describes an overstocked position. Green (G) represents an inventory position that requires no action. Yellow (Y) represents a part that has entered its rebuild zone. Red (R) represents a part that is in jeopardy. Dark red (DR, some authors refer to this as the black zone) represents a stockout. This color-coding system (the words used in the text and the abbreviations used in some diagrams as the figures are in black and white) will be used for both planning and execution priority and visibility and is integral to the power of the ASR solution. From a planning perspective, the color-coding will determine if additional supply is needed based upon the *available stock position (on hand + open supply - demand allocations [including qualified spikes])*. From an execution perspective, the color-coding will determine actions (primarily expediting or resource schedule manipulation) based on different types of alerts. This will be explained in the section titled “Highly Visible and Collaborative Execution.”

Because each part within a buffer profile has different individual traits, it yields individual buffer levels and

stratification zones for each part within that group (see [Fig. 12-4](#)). It is important to note that the zones need not be of equal proportions. Instead, the percentage of each zone is determined by the type of buffer profile to which the part belongs. The illustration to the right in [Fig. 12-4](#) shows three parts in the buffer profile group “A-10.” Each of the parts has a different top level and stratification levels because they have different individual part traits.

Note: companies will know if their buffer profiles are correct when the on-hand (not available stock) inventory position should average in the lower half of the yellow zone.

The color-coding also allows planners and executives to see how many overstocked as well as out-of-stock parts there are at any one time. If you combine the raw material value with the overstocked items, you can determine quickly how much excess cash is stored in excess inventory. Remember, that while being able to see stockouts is important what is really damaging is stockouts with demand allocations against them, which reinforces the need for visibility based on the available stock equation.

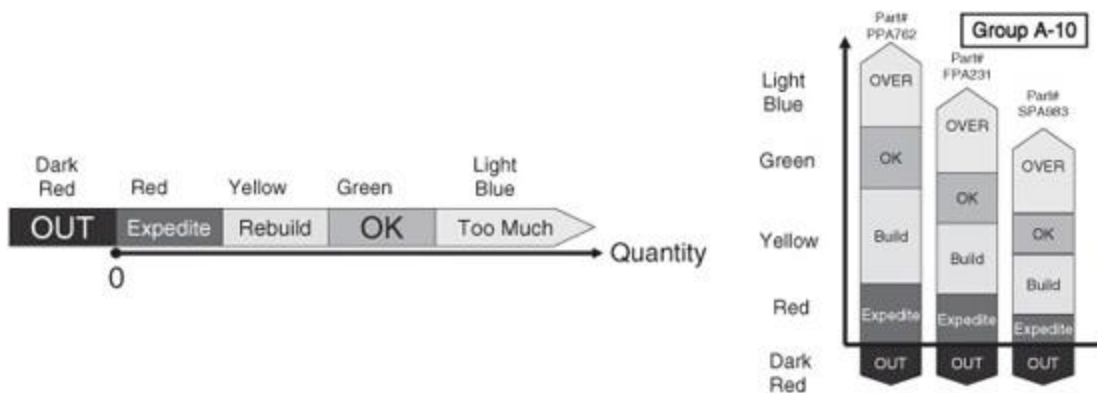


FIGURE 12-4 ASR stratifies the total buffer level into zones.

3. Dynamic Buffers

Over the course of time, group and individual traits can and will change as new suppliers and materials are used, new markets are opened and old markets deteriorate, and manufacturing capacity and methods change. Dynamic buffer levels allow the company to adapt buffers to group and individual part trait changes over a rolling time horizon. Thus, as these buffers encounter more or less variability, they adapt and change to fit the environment. Please note that the length of the rolling time horizon is very specific to the environment. Some companies may choose a 3-month roll while others must use 12 months. [Figure 12-5](#) illustrates how a buffer can adjust based on actual consumption. The initial buffer size (based on its buffer profile and individual part traits) can be seen at the far left of the figure. The black line represents the available stock position, while the grey line represents demand per week. Let us say that, for this part, we were using a three-month rolling time horizon. Over the course of a 24-month period, you can see that demand rose dramatically, began to taper off, and then eventually stabilized. The buffer followed the trend.

Additionally, these individual buffer profiles can be manipulated through something called “planned adjustments,” based on certain capacity, historical, and business intelligence factors. In ASR, these planned adjustments represent the necessary elements of planning and risk mitigation required to help resolve the conflict between predictability and agility. These planned adjustments are manipulations to the buffer equation that affect inventory positions by raising or lowering buffer levels and their corresponding zones at certain points in time. Planned adjustments are used for common situations like seasonality, product ramp-up and ramp-down, and capacity ramp-up and ramp-down. In the seasonality example in [Fig. 12-6](#), you can see a product that has a substantial bulge in demand once per year. In the part ramp-up example, you see a part that is being ramped up based on a sales and

marketing plan. In the part ramp-down example, you can see a part that is being discontinued. In all cases, if the planned adjustments do not follow the actual consumption (who ever heard of a Sales and Marketing Plan being accurate?), the color-coding/buffer stratification system will quickly identify that things are not going according to plan.

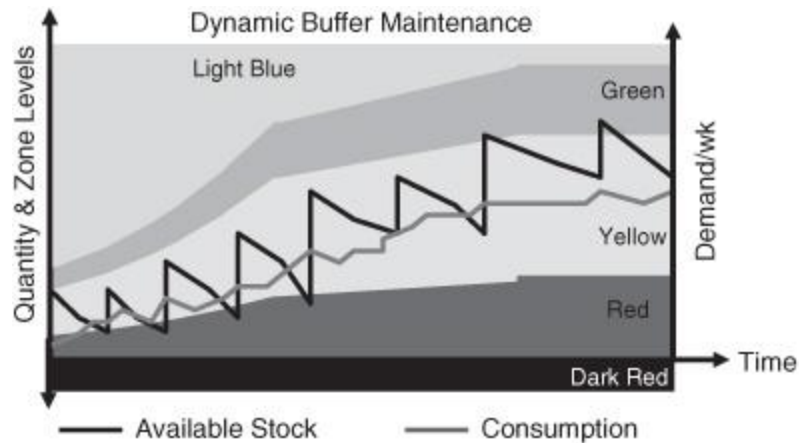


FIGURE 12-5 Dynamic buffer maintenance.

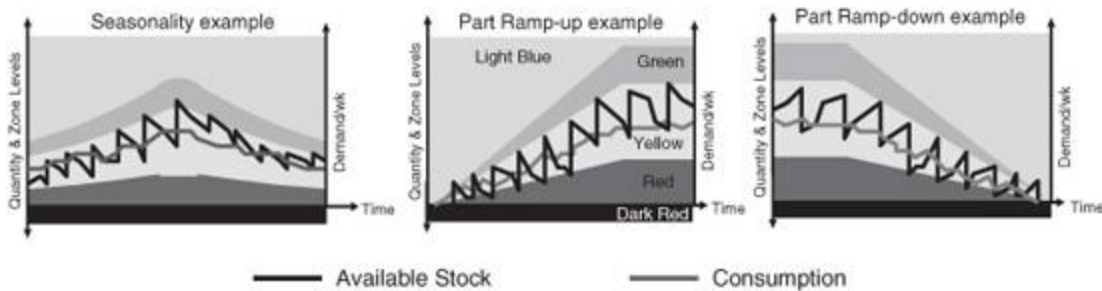


FIGURE 12-6 Buffer profile adjustments.

The combination of these first two solution elements (strategic inventory positioning and dynamic buffer level profiling and maintenance) of ASR creates strategically placed points of inventory that are actively managed, carefully sized, and dynamically adjusted. These buffers dampen or eliminate the effects of variation caused by the bullwhip effect and system nervousness that are passed up

and down the chain of resources and dependencies ([Fig. 12-7a](#)).

4. Pull-Based Demand Generation

Most Purchasing, Materials, and Fulfillment organizations have limited capacity and trust when it comes to sorting through the current demand signals and planned orders generated by MRP. The volume of reschedule messages is impossible to work before more changes happen and the process begins again. Many times critical actions are missed or incomplete pictures are painted. A significant understanding of MRP logic is required to even begin to understand the implications of a reschedule message. Many times it is easier to just leave it alone than risk disrupting the operation. However, this short-term fix can generate the need for many expensive corrective actions later (expedites, premium freight, overtime, etc.).

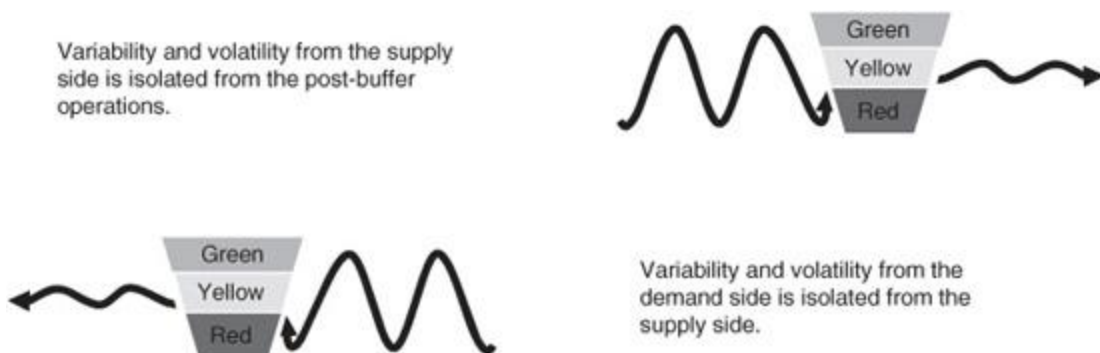


FIGURE 12-7a Conceptual illustration of dampening effect of stock buffers.

Generating, coordinating, and prioritizing all materials signals becomes much simpler when the environment is modeled properly. The current inventory status is evaluated for potential negative impacts and flagged for alert against open supply orders and demand allocations, which includes

future sales orders that meet specific spike criteria. Planners then have the ability to see where the signals are really coming from and react, before they get into trouble. This better matches the current intuition of the planners, but now they have the real visibility to establish correct and comprehensive priorities.

Key components of the ASR supply generation process include the following.

Demand Driven

Buffer levels are replenished as actual demand forces buffers into their respective rebuild zones. It is important to note that the buffer level driving demand generation is based on an “*available stock*” equation (as opposed to on-hand). *Available stock is calculated by taking on-hand inventory plus open supply minus demand allocations.* Actual on-hand inventory position relative to the buffer zones will provide the execution priority (discussed later in this chapter). [Figure 12-7b](#) shows the difference in relative buffer position between available stock and on-hand. The black arrows indicate the on-hand position and the white arrows represent the available stock position. This type of visibility gives relatively clear signals from both a planning and an execution perspective.

For example, part “f576,” according to its available stock position relative to its defined buffer zones, clearly needs additional supply created. Additionally, part “r672” does not need additional supply but rather the existing open supply needs to be considered for expedite.

Decoupled Explosion

Component part requirements are calculated by pegging down through the BOM. However, this planning is decoupled

at any buffered component part that is independently managed by an ASR buffer. This prevents the tsunami wave (nervousness) from rippling throughout the company as it does under MRP when a disruption occurs. The decoupled explosion from our earlier example of FPA is shown in [Fig. 12-8](#). Note that whenever a buffer position is encountered, the BOM explosion stops. The figure on the left depicts the explosion for the parent item FPA after its available stock position has been driven into the yellow zone. The middle figure represents buffered children that independently explode when they have reached their respective rebuild zones. Finally, you see the explosion for Sub-Assembly B (SAB) after its available stock equation has been driven into yellow.

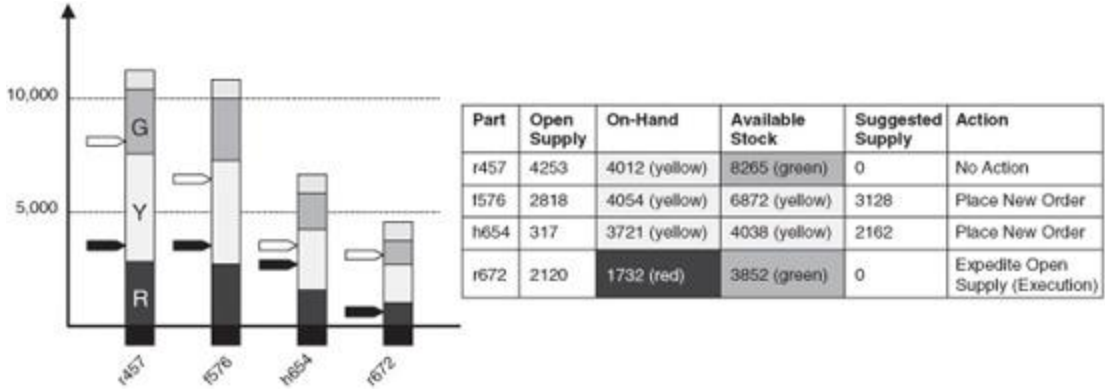


FIGURE 12-7b Available stock versus on-hand.

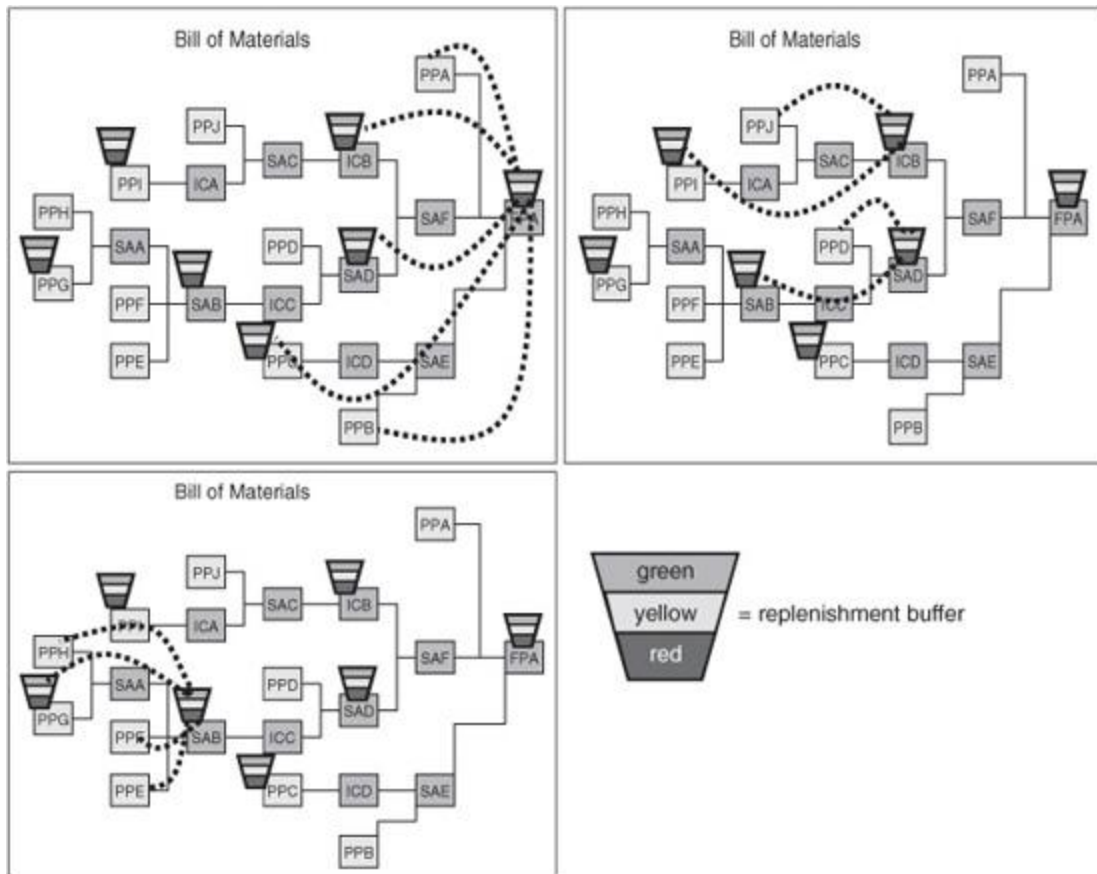


FIGURE 12-8 Decoupled explosion.

Material Synchronization

Component parts with incoming supply orders that are out of synch with demand allocations from parent work orders must be highlighted. This allows the Planners to take actions or make adjustments before work is released to the floor. This reduces the confusion in manufacturing and eliminates a significant amount of expediting.

ASR Lead Time

At the parent item level, MRP understands two types of lead times—neither of which are realistic for most environments.

The first is a fixed lead time called *manufacturing lead time*, which according to the *APICS Dictionary* (Blackstone, 2008, 78) is “*the total time required to manufacture an item, exclusive of lower level purchasing lead time.*” (© APICS 2008, used by permission, all rights reserved.) This is the most commonly used lead time definition in most MRP implementations. Believing the assumption that all parts will be available at the time of order release, however, is like sticking your head in the sand. Many MRP systems recognize another type of lead time called cumulative lead time. *Cumulative lead time* in the *APICS Dictionary* (Blackstone, 2008, 30) is “*the longest planned length of time to accomplish the activity in question. It is found by reviewing the lead time for each bill of material path below the item; whichever path adds up to the greatest amount of time defines cumulative lead time.*” (© APICS 2008, used by permission, all rights reserved.)

It is found by reviewing the lead time for each BOM path below the item; whichever path adds up to the greatest amount of time defines cumulative lead time. Most planners understand that the longer the cumulative lead time for the part, the more risk there is for disruption and volatility during that time or that the customer tolerance time will not allow for this lead time (analyzing these risks and breaking up the cumulative lead time is a key aspect of ASR’s Strategic Inventory Positioning solution element). With this cursory recognition, companies often hold intermediate or subassembly stock and stock on long lead time purchased items. These stock positions protect and compress lead times for end items. Simply put, this means that the realistic lead time for a manufactured part is neither the manufacturing lead time nor the cumulative lead time. In fact, the realistic lead time is determined by and defined as the longest cumulative unprotected leg in the BOM. This is called the *ASR lead time*. An illustration of the ASR lead time concept is shown in [Fig. 12-9](#). The first section of the figure

represents the BOM for a part called “20Z1.” Beside each unique part is a number that represents the manufacturing lead time of that part in days. As you can see, the manufacturing lead time for part 20Z1 is four days. The cumulative lead time path of 31 days is depicted in the second section of the figure as a bold line. The third section of the figure shows that when part 501P is buffered (depicted as grayed out) it shifts the longest unprotected sequence in the BOM to the bold marked path. The ASR lead time of 20Z1 is now 24 days, with the longest child contribution to lead time coming from 408P. Now, the fourth section of the figure shows that when we buffer 408P it shifts the ASR lead time to the other side of the BOM and is 21 days. The final section of the figure shows that this company chose to buffer the subcomponent 302. The ASR lead time for 20Z1 is now 11 days and the ASR lead time for part 302 is 17 days.

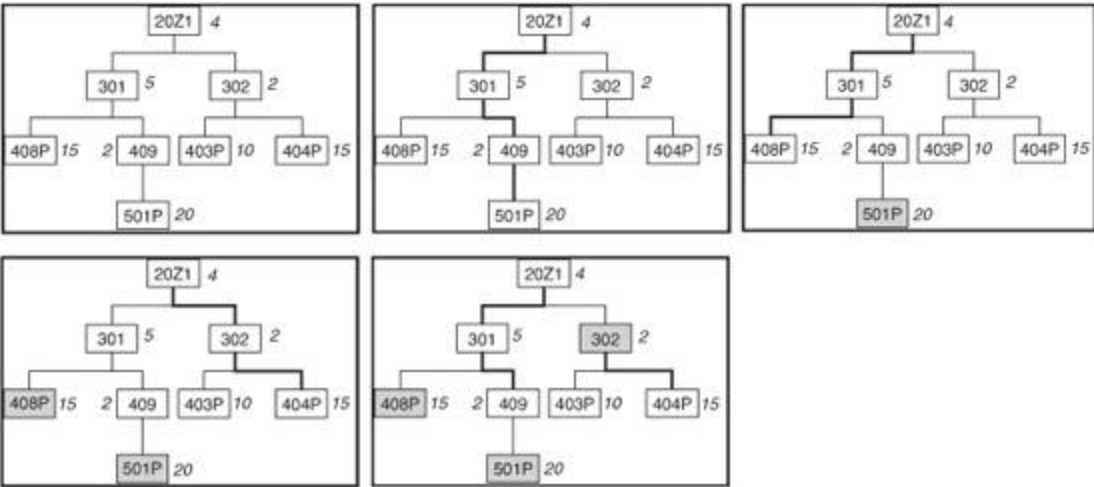


FIGURE 12-9 The ASR lead time concept.

Part #	Status Buffer Remaining	Available Stock	On-Hand	Open Supply	Demand	Recommended Order Quantity	ASR LT
FAE6721	OUT	-27	0	0	27	97	12
FAC6321	10%	100	142	32	74	900	17
BAC4321	42%	327	322	112	107	453	13
BAF6722	72%	89	47	63	21	0	4
BAE4322	112%	4625	4325	512	212	0	7

TABLE 12-3 ASR Planning Visibility

Highly Visible Priority

All ASR buffered parts are managed using highly visible zone indicators including the percentage of the depletion of the buffer (frequently called buffer penetration). This is a far simpler and faster approach than to have to sift through the planning queue checking all available stock equations to determine real priority. [Table 12-3](#) shows an example of ASR planning visibility. Again, the buffer status in ASR planning relates to the available stock position. The recommended order quantity will be the quantity to bring the available stock position to the top of the green (which is the top of the buffer).

Qualified Order Spike Protection

Most MRP systems force planners into a choice—bring in all known future demand or bring in none of it. The demand forecast consumption rules are some of the most complex areas to understand in even the most rudimentary MRP system. The big question is how to handle the overage or under consumption against the expected quantities. When MRP was planned in weekly time buckets, the choices were a bit easier, but now with MRP planning in daily quantities, the forecast error can be almost impossible to identify and

respond to in a timely fashion. The demand time fence will not allow the planner to realize that a qualified order in the planning horizon is looming and will cause enormous disruption to the plan once the order matures and crosses the demand time fence. In ASR, the buffer profiles and stratifications combined with the concept of ASR lead time allow qualified order spike protection over a realistic compensation time horizon. Thus, an order spike threshold is applied out over the ASR lead time to qualify sales orders that, according to the buffer profile, are spikes and will jeopardize the integrity of the buffer. This allows Planners to compensate effectively for known upcoming spikes in demand.

Realistic Lead Time Visibility

All orders are assigned due dates using ASR lead times. In an MTO environment, it is important to have ASR lead times visible because it can help focus any necessary expedite efforts and be used to make promises that are more realistic to customers. In *make-to-stock* (MTS) environments, ASR lead times are crucial because this is a more realistic parameter to help determine stocking levels as well as generate alert signals in execution.

In [Table 12-4](#) is a point-by-point comparison of the typical MRP implementation attributes that we detailed earlier versus applicable ASR attributes.

5. Highly Visible and Collaborative Execution

Simply launching *purchase orders* (POs) and *manufacturing orders* (MOs) from an ASR system's more effective pull-based planning mechanism does not end the materials management challenge. These POs and MOs have to be managed effectively to synchronize with the changes that

often occur within the “execution horizon.” The execution horizon is the time from which a PO or MO is opened until it is closed in the system of record.

	Typical MRP Attributes	Typical ASR Attributes	ASR Effects to the Organization
Planning Attributes	MRP uses a forecast or master production schedule as an input to calculate parent and component level part net requirements.	ASR uses known and planned part traits to set only the initial buffer size levels. These buffer sizes are dynamically resized based on real demand and variability. Buffer levels are replenished as actual demand forces buffers into their respective rebuild zones.	ASR eliminates the need for a detailed or complex forecast. Planned adjustments to buffer levels are used for known or planned events/circumstances.
	MRP pegs down the entire BOM to the lowest component part level whenever available stock is less than exploded demand.	Component part requirements are calculated by pegging down through the BOM. This planning, however, is decoupled at any buffered component part that is independently managed by an ASR buffer.	This prevents the tsunami wave from rippling throughout the company as it does under MRP when a disruption occurs. It eliminates system “nervousness.”
	MRP allows the release of work orders to the shop floor without consideration of component parts availability.	Projected available stock for component part requirements is verified prior to work order release to ensure work is not released to the floor if parts are not available.	Eliminates excess or idle WIP.
	Lead time for parent parts is the manufacturing lead time only for the parent, regardless of the cumulative lead time for parent and lower level component parts.	Lead time for parent parts recognizes both manufacturing lead time for the parent as well as the ASR lead time for non-buffered component parts on the longest unprotected leg of the BOM. Remember that the total cumulative lead time for the end item is decoupled at any strategic buffer points.	Creates a realistic lead time for customer promise and buffer sizing. Enables effective lead time compression activities by highlighting the longest unprotected path

	Typical MRP Attributes	Typical ASR Attributes	ASR Effects to the Organization
Stock Management Attributes	Fixed reorder quantity, order points, and safety stock that do not adjust to actual market demand or seasonality.	Buffer levels are dynamically adjusted as the part-specific traits change according to actual performance over a rolling time horizon.	ASR adapts to changes in actual demand.
	Only parts hitting minimum or reorder point are flagged for reorder.	All ASR buffered parts are managed using highly visible zone indicators including the percentage encroachment into the buffer. This gives you a general reference (color) and discrete reference (%).	Planning and Materials personnel are able to identify quickly which parts need attention and what the real-time priorities are.
	OPTION 1: Past due requirements and orders to replenish safety stock are often treated as “Due Now.”	All orders get an assigned realistic due date based on ASR lead times.	Creates a realistic lead time for customer promises and buffer sizing. Enables effective lead time compression activities by highlighting the longest unprotected path.
	OPTION 2: Or priority of orders is managed by due date.	All orders get an assigned realistic due date based on ASR lead times. All ASR buffered parts are managed using highly visible zone indicators including the percentage encroachment into the buffer. This gives you a general reference (color) and discrete reference (%).	Planning and Materials personnel are able to identify quickly which parts need attention and what the real-time priorities are.
	Limited future demand qualification. Limited early warning indicators of potential stockouts or demand spikes.	An Order Spike Horizon looks out over the ASR lead times of parts to identify large sales orders and qualify them as a spike in relation to the parts buffer levels. This allows the plan to compensate effectively for many spikes in demand.	Reduces the materials and capacity implications of large orders. Allows stock positions to be minimized since spike protection does not have to be “built in.”

TABLE 12-4 MRP versus ASR Attributes

ERP and MRP systems share the same “P” for planning. These are planning systems and not execution systems.

Most ERP and MRP systems lack real visibility to the actual priorities associated with the entire queue of POs, *transfer orders* (TOs), and MOs throughout the manufacturing operation and supply chain. Without this visibility, the supply chain (Suppliers, Manufacturing, Fulfillment, and Customers) employs the usual default mechanism of *priority by due date*.

Priority by due date often does not convey the real day-to-day inventory and materials priorities. Priorities are not static; they change as variability and volatility occur within the active life span of POs and MOs—the time from when they are opened until they are closed. Once again, this life span is called the “*execution horizon*.” Customers change their orders, quality challenges come up, there can be weather or customs-related obstacles, engineering changes happen, and suppliers’ capacity and reliability can temporarily fluctuate. The longer the execution horizon, the more volatile the changes are to priority and the more susceptible a company is to adverse material synchronization issues.

Ask yourself the following questions:

How does the manufacturing floor really know the relative priorities of stock orders?

- Does your operation ever have MOs to replenish stock that have the same due date (either a discreet date or “due now”)? How does the manufacturing floor decide what the priority is?
- Do you ever have MOs to replenish stock orders that have different due dates? Is it conceivable that despite an MO being due later, it is actually a higher priority based on certain events that have happened during the execution horizon? Have you ever built inventory in a rush only to find it sitting there for weeks while another MO could have averted a shortage if only you knew?

How does the supplier know how to align its capacity to your priorities?

- Do you ever have several open POs to a supplier all with the same due date? If yes, how do they know which is the most important to apply efforts to? If they call, can your Planner communicate the correct priority without having to research and peg or search for the source of various parts requirements up through the levels of the bills of materials. This is like searching through a spaghetti string mess.
- Do you ever have several open POs to a supplier with different due dates? Is it conceivable that despite a PO being due later, it is actually a higher priority, once again, due to changes that have occurred within the execution horizon? Have you ever paid for overnight charges only to find dust on that box months later?

Any sort of visibility to or a specific answer about the real-time priority of stock orders often necessitates a manual workaround or subsystem, which requires massive daily efforts of analysis and adjustments.

ASR Alerts

ASR provides real visibility of priority using a system of various types of alerts including:

- *Current Inventory Alerts* are for parts that are currently stocked out or in trouble. Is there demand for these parts or is the part just stocked out? There is a

difference in priority between parts that are stocked out with demand versus those that are stocked out without demand.

- *Projected Stockout Alerts* are for parts where projected consumption may result in a stockout prior to receipt of incoming supply orders. This is a radar screen that alerts materials and planning personnel to anticipated projected on-hand red zone penetrations over the ASR lead time of the part based on average daily usage and open supplies. If a company manages its Projected Stockout Alerts well, it will reduce the number of Current Inventory Alerts.
- *Material Synchronization Alerts* are for situations when any parts' demand and supply dates are out of synchronization. A part can have demand against it generated by a sales order or work order for a parent item. If the open supply promise date for this part is after the due date for the demand, then there is a potential negative available stock position. This means that demand and supply are not synchronized. This situation can occur when the demand moves in earlier than the open supply promise date or the open supply promise date moved out in time. Typically, this will drive either expediting action on the child component or the rescheduling of the parent item.
- *Lead Time Alerts* are used to prompt personnel to check up on the status of critical non-stocked parts before these parts become an issue (see the section titled "Lead Time Managed Components").

Visible Buffer Status

ASR allows actual order priorities (POs, TOs, or MOs) to be conveyed effectively without additional efforts,

disconnected subsystems, or other workarounds. Color-coding gives an easy to understand general reference. The percentage of buffer remaining gives a specific discrete reference. These references convey the actual priority regardless of due date. [Figure 12-10](#) shows examples of buffer displays for geographically distributed (by location), manufactured, and purchased items. Note how the due date may not correspond with actual priority (WO 819-87). Additionally, observe on the Purchased Items display how easy it is to identify priority when things are due on the same date.

Lead Time Managed Components

Many critical components simply do not make sense to stock due to their relatively low volume. Ask most seasoned materials managers in major manufacturers and they can immediately recite a list of these types of components. These long lead time components can be very difficult to manage especially if they are sourced from a remote supplier. Without an effective way to manage these parts, we risk major synchronization problems, costly expediting, and poor service level performance. In ERP/MRP systems there is very little done about the management of these parts. They are managed by due date with no formal system of visibility and proactive management to reflect real priorities. The assumption remains as it did when MRP was first developed that all the parts will be available at the time of the order that needs them being released. Only when those parts are missing do personnel become aware of it and then expediting begins. The problem is only identified when the part is late. Orders using that part are released short those parts, causing possible rework on the shop floor and increasing work in process. Alternatively, some companies will begin to pull parts ahead of time to identify

this kind of shortage. This process results in a storehouse of partially filled kits and a manual system to track the missing parts.

<i>Purchased Items</i>			<i>Manufactured Items</i>				<i>Distributed Items</i>		
Order #	Due Date	Buffer Status	Order #	Due Date	Item #	Buffer Status	Item #	Location	Buffer Status
PO 820-89	05/12/09	Critical 13%	WO 819-87	05/24/09	GADC843	Critical 13%	GADC843	Region 1	Critical 11%
PO 891-84	05/12/09	Med 39%	WO 832-41	05/22/09	GCDC632	Critical 17%	GADC843	Region 2	Med 41%
PO 276-54	05/12/09	Med 41%	WO 211-72	05/22/09	FCDG672	Med 34%	GADC843	Region 3	Med 36%

FIGURE 12-10 ASR buffer displays for geographically distributed items.

ASR gives special status and visibility to these parts. These lead time managed components are tracked and at a defined point in the part’s lead time, buyers are prompted for follow-up. If satisfactory resolution is not achieved, the visible warning or alert continues to rise in priority. Resolution could be either the assignment of a follow-up date (temporary resolution) or the assignment of a final confirmed date and decision (could be sooner, on time, or later). Regardless of what the resolution is, at least it is known and understood ahead of time. Then the other parts affected can be reprioritized. Additionally, these types of proactive efforts often nip potential problems in the bud, resulting in better due date performance for these types of components.

The purpose of the ASR execution concepts is to increase the amount of accurate and timely information available to the entire chain. This highly visible and collaborative execution capability creates a remarkably effective supply chain that can respond to real market demand without manual workarounds and other disconnected subsystems. Purchasing, Manufacturing, and Fulfillment personnel thus are able to see and communicate a bigger picture that is clear, concise, prioritized for action, and shows the ramifications of decisions and actions based on the

dependencies in the aggregate material supply and fulfillment system.

The five ASR components (strategic inventory positioning, dynamic buffer level profiling and maintenance, dynamic buffers, pull-based demand generation, and highly visible and collaborative execution) work together to dampen the nervousness of MRP systems and the bullwhip effect on MRP systems in complex and challenging environments. Utilizing this ASR approach, the planners no longer must try to respond to every single message for every single part that is off by even one day. The ASR approach provides real information about those parts that are truly at risk of negatively affecting the planned availability of inventory. ASR sorts the significant few items that require attention from the magnificent many parts that are being managed. Under the ASR approach, fewer planners can make better decisions more quickly.

ASR Implementation Considerations

1. What happens to inventory levels in an ASR implementation? Similar to the focus on Lean manufacturing, while significant inventory reductions are an effect of implementing the ASR approach, this concept is not intended to focus on inventory reduction. Inventory is a result rather than a focus. ASR should never be implemented with the sole purpose of inventory reduction. Dramatic reductions in inventory, however, are a result of the overall approach rather than the primary objective. The system drains the inventory that is not needed for real protection of due date performance. Now the inventory that is in the system is really working and generating a positive *return on investment* (ROI).

In early adoptions of the ASR approach, the impact on inventory is consistently somewhere between a 20 to 50

percent reduction in the first year. However, at the earliest stages of the implementation there is typically a temporary increase in overall inventory levels because parts may need to be buffered that were not previously inventoried. This additional inventory is combined with substantial inventory dollars that exist over the top of the required ASR buffers. As that excess inventory (items now residing in the blue zone) drains down to within the buffer parameters, then companies begin to see significant inventory reduction and a highly improved level of turns.

2. Does my ERP system offer ASR functionality? At the time of this writing, no ERP system has the total functionality identified in this chapter. Most systems do support both min/max as well as MRP with an input of a forecast or master production schedule (MPS) for inventory planning. None of this push-based approach really enables the five components of ASR. Min/max levels are static and usually are not reviewed after the initial system setup. There is no adaptive mechanism to update levels based on the experienced volatility and actual demand on the system. Remember that forecasting and MPS are inherently to drive a push system. ASR is inherently a demand-pull and replenishment system—sensing and adapting to actual market demand and volatility. The BOM decoupling analysis is not supported by any ERP system today. This decoupling is a key ASR component for positioning the break walls to absorb cumulative variability arising from both supply and demand. The concept of ASR lead time is totally foreign to any ERP system. ASR lead time is crucial in understanding where to compress and manage lead times within a BOM through the use of decoupling and to determine the proper stocking/buffer levels of a part.

3. What are the specific business benefits expected from implementing ASR? In addition to resolving the MRP compromises and the effects associated with them, there

are additional business benefits when the ASR approach is implemented.

- a. **Protect and increase flow** by significantly reducing the negative impact of variability in dependent and interdependent systems. This can include both demand variability from the marketplace and supply variability starting with external sources and then continuing internally through operations.
- b. **Create a decisive competitive advantage** by developing and exploiting ways to significantly compress product and materials lead times to the marketplace. This ensures that lead time offers are significantly better than what the market is expecting. In most cases, a highly competitive lead time can be achieved with no investment in equipment or traditional lead time reduction initiatives.
- c. **Highly improved on-time delivery performance** to the marketplace. If lead times are dramatically reduced and flow is improved, then significant improvements in service performance can and will follow. This provides another opportunity for a competitive advantage in the marketplace.
- d. **“Right size” inventory** through the strategic inventory positioning process. This ensures that the right amount of protection is carried in the right places based on the rate of demand pull from the market and potential disruptions in supply and demand. The critical difference with ASR is that these are dynamic buffers that constantly reflect the changing market and supply conditions.
- e. **Enable better execution.** The ongoing management process in ASR becomes relatively simple once the analysis is complete and buffers are established in the correct places. The execution side ensures early identification of potential problem areas such as a supplier that is going to be late or a delayed

work order that could potentially impact buffers. This allows action to be taken before these small disruptions become big problems.

4. What kind of manufacturing environments should consider ASR? Characteristics of environments where ASR delivers the significant business benefits listed are as follows. The more of these characteristics that an environment has, the more significant the benefits will be.

- Environments with sets of highly repetitive builds (either product or process).
- Environments that will reward you for shorter lead times through either premiums or increased sales.
- Environments that frequently use the same purchased component or raw items.
- Environments that utilize the same components across multiple parent parts.
- Environments with deep and complex BOMs.
- Environments with longer or more complicated routings that create significant scheduling or lead time difficulties.
- Environments that are considering or currently using pull-based scheduling and execution.

Case Studies

In early implementations of this approach, a very powerful insight was realized—*the business benefits are complementary and happen collectively*. Unlike the typical expectation of inventory versus customer service tradeoffs, in the early implementations of ASR there have been no

tradeoffs. Not only does inventory significantly go down, customer service dramatically improves.

Case Study 1: Oregon Freeze Dry

Oregon Freeze Dry is the world's largest custom freeze dryer. Prior to implementing ASR, they used traditional MRP with standard minimum batch practices. By implementing ASR only (no DBR or S-DBR) with no additional capital expenditure, overhead, or other improvement initiatives, Oregon Freeze Dry reported the following gains: Mountain House Division:

- Sales increased 20 percent
- Customer fill rate improved from 79 to 99.6 percent
- This was accomplished with a 60 percent reduction in inventory

Industrial Ingredient Division:

- 60 percent reduction in MTO lead time
- 100 percent on-time delivery
- This was accomplished with a 20 percent reduction in inventory

Case Study 2: LeTourneau Technologies, Inc.

The LeTourneau Technologies, Inc.™ (LTI) companies include some of the world's leading innovators in manufacturing, design, and implementation of systems and equipment for

mining, oil and gas drilling, offshore, power control and distribution, and forestry. LTI has two main manufacturing facilities (Longview, TX and Houston, TX) that are similar in terms of capability, product complexity, and size. One can see the dramatic differences between the two comparable campuses of Longview and Houston in the following information. To be very clear, the type of manufacturing is very similar in terms of both complexity and scale.

Beginning in 2005, the market began to take off for all LTI business segments. What is important to understand is that LTI has been through these boom cycles before. All previous times, however, LTI’s inventory and expenses have dramatically risen at a similar rate as revenue along with deteriorating service levels. What is unique about this particular case is that the Longview facility using ASR (as well as a partial implementation of DBR) was able to dramatically control inventory and expenses while maintaining excellent service levels in the boom cycle.

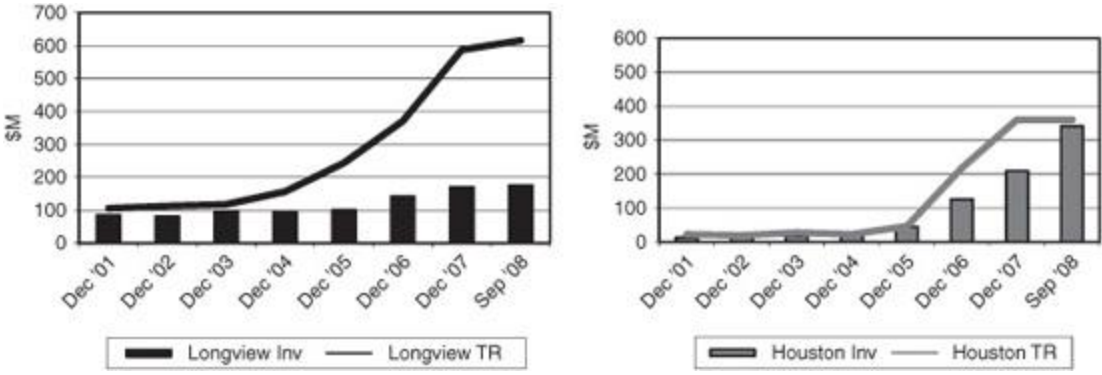


FIGURE 12-11 Total revenue versus inventory.

Additionally, what should be noted is that all boom markets eventually end. One can see in the graphs in [Fig. 12-11](#) that in 2008 the markets began to cool off. When those boom times are over, ASR minimizes exposure to inventory liabilities. The bottom line is that no matter what kind of economic times a company finds itself in, good

inventory practices that minimize inventory exposure while maintaining service levels is always the right strategy.

This first graph in [Fig. 12-11](#) shows Total Revenue versus Inventory from 2001 to 2008 from the Longview campus only. Note that beginning in 2005 there was rampant growth. Revenue grew by a factor of greater than 300 percent (over \$400 million). Over that same period, inventory rose only by 80 percent (approximately \$80 million).

This second graph in [Fig. 12-11](#) shows Total Revenue versus Inventory from 2001 to 2008 from the Houston campus only. Note that at the beginning of 2005 there was the same rampant growth curve as observed in Longview. In this case, however, inventory ended up growing at nearly the same rate as revenue. There is about a six to nine month lag, but it is pacing at the same rate. Why is there a lag? As is typical with most MRP implementations, the plant is building to forecast.

Now, as can be seen in both graphs, when the market begins to turn at the beginning of 2008, LTI is exposed with a huge amount of inventory liability. In fact, due to the nature of forecasting and long lead times, there is a risk that the inventory will actually grow beyond revenue in the short run without massive course correction in the form of PO and MO cancellation or delay. This is a classic effect of traditional MRP driven environments.

It is very important to note that the people in the Houston facility are smart, professional manufacturing personnel. They simply did not have the tools and new approaches at their disposal to replicate what happened at Longview. The graph in [Fig. 12-11](#) is not an indictment of those people; it is proof that traditional MRP represents a huge liability in the volatile and variable manufacturing environments that tend to be today's rule rather than exception.

For more on LTI's case study, see [Chapter 14](#).

Summary

By bringing together rules, vision, and technology, ASR provides a practical real world solution to the MRP conflict found in so many companies today. ASR still allows the company to utilize its formal planning system and finally realize the ROI expected when the system was first implemented. The current ERP system is not ripped out and replaced. Instead, the components of ASR leverage all the good work done to date. The five components of the ASR approach effectively manage the volatility and variability plaguing your company to create the velocity and visibility necessary to provide a competitive advantage in today's hypercompetitive market. Isn't that better than disconnected sticky notes and Excel spreadsheets?

The authors have set up the Web site www.beyondmrp.com for interested readers to learn more about ASR. We welcome your thoughts and feedback on this innovative approach.

References

- Aberdeen Group. 2006. *The ERP in Manufacturing Benchmark Report*. Boston: Aberdeen Group.
- Blackstone Jr., J. H. 2008. *APICS Dictionary*. 12th ed. Alexandria, VA: APICS.
- Masson, C., Smith, A., and Jacobson, S. 2007. "Demand driven manufacturing," AMR Research Report. January 2007, AMR-R-20105.

Orlicky, J. 1975. *Material Requirements Planning*. New York: McGraw-Hill.

About the Authors

Chad Smith is the cofounder and Managing Partner of Constraints Management Group (CMG), a services and technology company specializing in pull-based manufacturing, materials, and project management systems for mid-range and large manufacturers. Mr. Smith has a wide range of experience in successfully applying pull-based systems within a diverse scope of organizations and industries. Clients, past and present, include LeTourneau Technologies, Boeing, Intel, Erickson Air-Crane, Siemens, IBM, The Charles Machine Works (Ditch Witch), and Oregon Freeze Dry.

Since the late 1990s, Mr. Smith and his partners at CMG have been at the forefront of developing and articulating the concepts behind ASR as well as building ASR compliant technology. CMG's homepage is at www.thoughtwarepeople.com.

Additionally, Mr. Smith is an internationally recognized expert in the application and development of TOC, getting his formal training at the Avraham Y. Goldratt Institute Academy and working under the tutelage of Dr. Eli Goldratt, author of *The Goal*, for several years.

Contact him at **csmith@thoughtwarepeople.com**.

Carol Ptak is at Pacific Lutheran University as Visiting Professor and Executive in Residence after years of executive management experience at PeopleSoft and IBM Corporation as well as many years of consulting expertise. Most recently, Ms. Ptak served as the Vice President and Global Industry Executive for Manufacturing and Distribution industries at PeopleSoft.

Here she developed the concept of Demand Driven Manufacturing (DDM) as an overall product and marketing strategy to align product development, market awareness, and demand generation. Her innovative approach is credited with significantly improving the company's position in the manufacturing industry software market and earned her national recognition in publications such as *CFO Magazine* and the *New York Times*. Prior to her accomplished record at PeopleSoft, Ms. Ptak spent four years at IBM Corporation starting as a member of the worldwide ERP/SCM solutions sales team and quickly rising to the position of global SMB segment executive.

From 1993 through 1998, Ms. Ptak owned Eagle Enterprises, a consulting firm that promotes company-wide excellence through education and successful implementation. She worked with a wide-range of businesses including internationally known corporations such as Boeing and Starbucks.

Ms. Ptak, who holds her MBA from Rochester Institute of Technology, is also certified by the American Production and Inventory Control Society in Production and Inventory Management (Fellow level) and in Integrated Resource Management and completed additional graduate work at Stanford University. She holds a bachelors degree in Biology from the State University of New York at Buffalo. Contact her at **ptakca@plu.edu**.

SECTION IV

Performance Measures

CHAPTER 13

Traditional Measures in Finance and Accounting, Problems, Literature Review, and TOC Measures

CHAPTER 14

Resolving Measurement/ Performance Dilemmas

CHAPTER 15

Continuous Improvement and Auditing

CHAPTER 16

Holistic TOC Implementation Case Studies

Problems with traditional accounting methods and other traditional measures are considered here. New methods for “Throughput Accounting” are discussed in terms of improving organizational performance. This method emphasizes financial measures that focus on global performance of the organization in contrast with measures that emphasize local/silo measurements. The erroneous traditional assumption that local optima accumulate to overall improvements in system performance is effectively challenged. In this context, the shortcomings of traditional cost accounting are discussed in depth.

Basic measures and processes of ongoing improvement that focus attention on indentifying local actions resulting in organizational improvement are presented. Elements

essential for other measures encountered in production operations, projects, and services such as buffer management, quality measures, service response times, and the like are presented along with hands-on solutions for structuring and implementing them. Traditional measures create actions in one function or department that cause conflict with other functions or departments. Cross-organization conflicts that can be created by measures and the resolution of these conflicts are also discussed.

The last two chapters describe the Process of Ongoing Improvement (POOGI) and the requisite auditing function and provide two detailed holistic case studies. Achieving POOGI in any organization not only requires a reliable focusing mechanism (to identify where and what to change and when and what not to change), but also a holistic decision support mechanism (to judge the system-wide or global impact of changes). Then, a fast and reliable feedback mechanism is needed for auditing progress/compliance or for identifying other important system performance gaps or variations.

CHAPTER 13

Traditional Measures in Finance and Accounting, Problems, Literature Review, and TOC Measures

Charlene Spoede Budd

Introduction

This chapter is a basic introduction to *Throughput Accounting* (TA). To provide historical perspective, the chapter provides a brief review of both the business environment and the development of cost accounting methodologies.

Accounting personnel usually are among the last people to be educated in *Theory of Constraints* (TOC) concepts. We are constantly amazed at the reported successful TOC implementations that have not educated accounting and finance people at all. Yet operations people expect that they can overcome resistance to their improvement plans. One very successful TOC implementation champion, strongly supported by the CEO, lamented that he could not understand why the accounting department had hired additional personnel to track the cost of each individual operation when he had established a seamless flow line. The accountants were doing what they had been taught to do. Without an understanding of TOC concepts, when they had sufficient information, they would begin questioning the

cost of certain operations, reporting local efficiencies, and providing other misleading information such as unit costs.

I hope that this chapter will develop an appreciation of accounting and finance personnel—what they can do for you and what they can do against you—and provide a strong argument for educating accounting and finance personnel along with those in operations. TOC initiatives need collaborative partners rather than colleagues who constantly construct a maze of barriers. For accountants who have a suspicion that traditional accounting methodologies produce internal data containing weaknesses for decision makers, this chapter will point out where the weaknesses exist. To accomplish this ambitious goal, the chapter will:

Copyright © 2010 by Charlene Spoede Budd.

- Briefly describe the history of traditional cost accounting and explain why it no longer provides the information needed to support the improvement initiatives made possible by TOC.
- Survey, classify, and describe the limitations of the profession's various *accepted* (published) solutions to replace traditional internal measurement systems.
- Discuss the breadth of TA and its impact on all TOC initiatives through a continuing case study.
- Identify the need for future research in accounting and finance to support TOC concepts, including the development of relevant performance evaluation systems.

The final section of the chapter will introduce the remaining chapters in this section of the *TOC Handbook*.

Traditional Cost Accounting and Business Environment

Cost accounting is designed and developed to help managers make decisions. When cost accounting's assumptions mirror those of the organization, the information provided enables good decisions. Conversely, when accounting assumptions are not valid, managers make good decisions only by using their intuition or by chance and not by using the accounting information provided.

As the environment changes, internal accounting and reporting should be changed to reflect that new environment and provide information that is more relevant to managers. In most companies, as we shall see, this adaptation has not occurred or has greatly lagged changes.

Development of Cost Accounting

Accounting has been around since exchanges first began taking place, but until the 19th century, few people were involved in financial reckonings and internal accounting was mostly conducted visually by owners and managers. With the onset of the industrial revolution and the growth of large companies, accounting became more important and cost accounting began to be developed to control large organization chaos (Kaplan, 1984; Cooper, 2000; Antonelli et al., 2006; McLean, 2006).

Since the industrial revolution began first in Great Britain, their engineers and accountants were the first to recognize the need for cost/management accounting (Fleischman and Parker, 1997; McLean, 2006; Edwards and Boyns, 2009), but their accounting developments were not publicized (Fleischman and Parker, 1997). In the United States, modern cost/management accounting began in the late 19th and

early 20th centuries (Tyson, 1993), especially with the introduction of mass production.¹ The scientific management movement, supported by the theories of Frederick Taylor² (1911, 1967), Walter Shewhart (1931, 1980), and Mary Parker Follet's enlightened approach to management (Follett and Sheldon, 2003), drove the development of supporting cost/management accounting by engineers and accountants such as Alexander Hamilton Church (Litterer, 1961; Jelinek, 1980), who railed against "averaging" production overhead costs over all jobs or products produced and insisted that all production costs must be assigned to orders or products (Church, 1908). In a two-stage process, overhead typically is assigned first to departments and then to jobs or products passing through the department.

Business Environment, First Half of the 20th Century

Frederick Taylor's theories, while not universally accepted, were widely believed and practiced by companies during the early decades of the 20th century (Kanigel, 1997). Business schools, including Harvard Business School (Cruikshank, 1987) began teaching Taylor's scientific management.

By the 1930s, most large manufacturers had adopted some form of manufacturing overhead allocation, but standard costs and related detailed variance analysis did not come into widespread use until after World War II (Johnson and Kaplan, 1987). Rather than being developed to control manufacturing costs, the original purpose of variance analysis was to value inventories and derive income statement costs (Johnson and Kaplan, 1987). This is because *generally accepted accounting principles* (GAAP) require that actual (not standard) costs appear on the balance sheet and the income statement and standard

costs, plus favorable variances or minus unfavorable variances, equal actual costs.

During World War II, the demand for war supplies fueled widespread implementation of mass production (Grudens, 1997). Following the war, companies rushed to fulfill pent-up consumer demand, and some companies used standard costs and variance analysis to control production costs (McFarland, 1950; Vangermeersch and Schwarzback, 2005).

Business Environment, Second Half of the 20th Century

During the 30 years following World War II, U.S. companies basically followed the same strategy of cost-conscious mass production. In addition, the *premier* department in schools of business, those drawing the most intelligent students and wielding the most political power, slowly switched from accounting to finance.

During the 1960s and probably much earlier, Harvard Business School began teaching MBA students how to *manage by the numbers*, meaning using a company's financial records and other formulas and models developed in finance (Peters and Waterman, 1982, 30) to manage the company. While there were cautions published about the formulas' complex and fragile treatment of uncertainty in the development of financial models and the overreliance on the skills of MBAs (Hayes and Abernathy, 1980, 67; Peters and Waterman, 1982, 31-33; Johnson and Kaplan, 1987, 15, 125-126), the predominance of finance departments over accounting departments in both academia and industry gradually spread across the United States and throughout the world during almost the next two decades.³ The focus of the 1980s on behavioral consequences of the formulistic approach to business decisions lasted about a decade; by the 1990s, though,

business was back to using formulas and models, along with new continuous improvement methodologies (Dearlove and Crainer) in order to regain ground lost to international competitors.⁴ This movement no doubt was inspired and certainly facilitated by consulting firms who found ingenious ways to convince management that their assistance was required (Stewart, 2009).

Accounting's Response to a 20th Century Changing Environment

Management accounting, for most companies, barely noticed the changes occurring in business due to their general absorption with other accounting areas (e.g., financial, tax, auditing) and most especially GAAP accounting for external reporting (Johnson and Kaplan, 1987, 2-14, 125). However, it became obvious during the 1980s that traditional accounting and accounting reports had lost their relevance for internal decisions (Johnson and Kaplan, 1987).⁵ *Fully absorbed* manufacturing costs, including variable and fixed costs of production (whether actual or standard costs), accumulated for external reporting purposes typically do not provide information needed by managers for operating decisions.

Some solutions to the irrelevance of management accounting information have been known for a number of years, but have not been widely accepted and practiced. In addition, newer solutions recently have been proposed (Kaplan and Norton, 1992; Johnson and Broms, 2000; Smith, 2000; Cunningham and Fiume, 2003; Oliver, 2004; Van Veen-Dirks and Molenaar, 2009). The most well-known proposed accounting solutions are discussed in the following sections.

Direct or Variable Costing Income Statement

Direct or variable costing⁶ (where all costs are divided into fixed and variable components that are then recorded in separate accounts), however, has been included in textbooks since at least the 1960s (Dopuch and Birnberg, 1969, [Chapter 15](#)) and has been covered in virtually every cost and management accounting textbook since that time (Hilton, 2009, [Chapter 8](#); Garrison, Noreen, and Brewer, 2010, [Chapter 7](#)). The basic format begins with revenues earned, then subtracts all variable costs to provide contribution margin (sometimes called gross margin). From contribution margin, all fixed costs (both manufacturing and general, selling, and administrative) are deducted to arrive at operating income. This method is not acceptable for external financial statements, however, and has not been broadly accepted.

Direct or variable costing is presented in all cost and managerial accounting textbooks as a method of periodic reporting and for providing information for decision makers. The basic idea is that revenue, minus all variable costs (basically equivalent to out-of-pocket costs), is subtotaled as contribution margin. Fixed costs, that must be incurred each period, are then subtracted from contribution margin to find operating income. While contribution margin may be used to find a contribution margin per constraining unit (discussed later), this topic is treated independent of the direct costing income statement. In addition, many accountants were taught that *direct labor*, the cost of workers actually transforming a company's product ("hands-on" work) is a variable cost, as was true when cost accounting was developed at the start of the 20th century.

Because TA follows the same basic format as direct or variable costing accounting for periodic reporting, however, this discussion is important.

Advantages of Direct Costing

Advocates for direct costing base their interest on internal flows (Dopuch and Birnberg, 1969, [Chapter 15](#))⁷ and providing information for internal decisions. They claim that direct costing:

- Focuses attention on those costs that most closely approximate the marginal (incremental) costs of production;
- Relates profit to sales, rather than to sales and production, as does traditional accounting;
- Treats fixed costs as a period expense since these amounts must be expended in order to be in position to produce and must be incurred each period without regard to production quantity (that is, certain costs must be incurred even when production is at or near zero).

The advantages and disadvantages of direct costing are discussed in every cost or management accounting textbook when the methodology is introduced. While supporters and detractors at one time were passionate in their support or opposition, most authors now just list the advantages and disadvantages.

Disadvantages of Direct Costing

Opponents of direct costing do not accept the benefits claimed for direct costing and point out its theoretical weaknesses. They claim, as support for their stance, that direct costing:

- Violates the matching concept of accounting where the total unit cost of production (variable and fixed) must be

recognized in the period when units are sold and not match the period in which they were incurred;

- Does not account for the total costs of producing a product on a recurring basis and that full (totally absorbed) costing, as in traditional accounting, is a better measure of the incremental cost of production;
- Is only applicable over a specified output range and variable costs may change outside the originally assumed range. (Of course, fixed costs are subject to the same claim.)

While it is difficult to know what proportion of companies regularly record variable and fixed costs in separate journal accounts (companies are not required to disclose this information), the author's experience is that most companies do not. Since the separation of costs into fixed and variable components is required for virtually all internal decisions, this information must be accumulated in special studies. Unfortunately, most accounting/finance departments have little time to devote to special projects when relevant information is not readily available.

Activity-Based Cost Accounting

Activity-based cost (ABC) accounting might be considered accounting's attempt to "return to the basics" with several new twists. First, overhead is assigned to many pools, not to departments. Second, since all overhead costs can be changed over time, they are assumed to be variable. Third, the selection of allocation bases (*drivers*) used to allocate pool costs depends on whether costs are incurred at the unit, batch, product line, or facility level. (Facility-level costs include general manufacturing costs common to all production.)

In practice, companies implementing ABC accounting allocate manufacturing costs as originally proposed by Church (1908, 1917), but rather than assigning overhead costs to departments, they are allocated to activity *pools*. Activity pools are holding accounts where costs for a particular activity, such as material movement, can be accumulated prior to being charged to users of the activity. Thus, if one product or product line requires more movement than other products, it would receive more of the material movement costs (Cooper et al., 1992).

Advantages of ABC Accounting/Management

Besides appearing more precise than traditional accounting, ABC accounting offered several advantages. ABC accounting:

- Gave companies (and accounting departments) hope that they could do something to reverse their poor business performance;
- Validated claims by operations people that small runs of “special” products cost more to produce than did long runs of common or commodity-type products;
- Silenced, for the most part, “fairness” arguments concerning overhead allocations;
- Provided much detailed information that could be analyzed for improvement initiatives, leading to the development of *activity-based management* (ABM; Cokins, 2001).

In addition to these advantages, companies that have taken the first step of charting all flows from purchase of raw materials through production processes to finish goods

and shipping prior to implementing ABC have universally reported benefits achieved from their increased knowledge of their systems. Making use of all the data collected and updating it to track frequent changes in the business environment, however, has proven quite difficult and costly.

Disadvantages of ABC Accounting

As originally developed, ABC accounting and ABM required tremendous amounts of quantitative data on anticipated and actual driver (allocation base) consumption.⁸ Complex original implementation efforts and continuing data collection issues resulted in complaints that ABC accounting:

- Requires too much detailed data collection efforts by operating employees who did not want or use the information provided;
- Permits subjective selection of pools and drivers;
- Lacks an easy way to identify erroneously reported data;
- Mixes fixed and variable costs in the same pool (by assuming all costs are variable);
- Focuses on reducing costs, not generating revenue;
- Generates costs that greatly exceed benefits attained (Palmer and Vied, 1998; Geri and Ronen, 2005; Bragg, 2007a, 207-209; Ricketts, 2008, 54.).

Even though the adoption of ABC accounting or ABM has been low and scattered (Kiani and Sangeladji, 2003; Cohen et al., 2005; Bhimani et al., 2007), academics and consultants continue to support the methodology (Stratton et al., 2009).

Balanced Scorecard

Recognizing the importance of appropriate performance measures to motivate employees to coordinate their activities (and later to implement company strategy), a performance scorecard that included nonfinancial metrics was developed by industry leaders.⁹ “The scorecard measures organizational performance across four balanced perspectives: financial, customers, internal business processes, and learning and growth.” (Kaplan and Norton, 1992; Kaplan and Norton, 1996, 2.) While the most well-known advocate of the balanced scorecard has not abandoned activity-based costing (Kaplan and Norton, 1996, 55-57), a more *balanced* set of measures is intended to include the guidance of nonfinancial metrics and provide a more global perspective.

Surveys indicate that balanced scorecard concepts are used in most large companies in the United States and throughout the world. Despite reported successful implementations, however, there is little empirical evidence that implementing a balanced scorecard results in increased earnings (Speckbacher et al., 2003).

Advantages of a Balanced Scorecard Performance Reporting System

One of the major benefits of a balanced scorecard system, its ease of understanding, also may be one of its biggest flaws. It is entirely possible that managers, easily accepting the basic idea of balancing metrics across all aspects of a business and therefore prone to implementing balanced scorecards without outside consultants, have not sufficiently customized their balanced scorecard systems. Nevertheless, companies expect that a balanced scorecard performance reporting system will:

- Focus all employees on longer-term goals;

- Clarify the relationships and importance of various strategic goals;
- Align employee behavior with strategic goals;
- Provide relevant and timely feedback to managers;
- Promote better decisions;
- Improve operating performance (Lawson et al., 2003; Buhovac and Slapnicar, 2007; Anonymous, 2008, 80).

Unfortunately, most of these expectations are unfulfilled for the majority of adopters.

Disadvantages of Balanced Scorecards

It is estimated that up to 70 percent of organizations have adopted *balanced* scorecards (Angel and Rampersad, 2005). However, even proponents of balanced scorecards admit that up to 90 percent of adopters fail to execute well-planned strategies (Weil, 2007). For whatever reasons,¹⁰ balanced scorecard promises have not been delivered. One author came up with a list of “Top 10” problems with most scorecards (Brown, 2007, 9). Most researchers conclude that a balanced scorecard:

- Encourages too many measures that divert focus from what is important;
- Gives obvious priority to financial measures; bonuses are rarely based on non-financial metrics;
- Excludes, too often, appropriate measures for learning and organizational growth;
- Provides an unfavorable cost/benefit ratio;
- Produces measures from diversified divisions that cannot be aggregated at the corporate level;

- Neglects to clearly connect strategy with action at the individual employee level;
- Provides lagging metrics that do not produce timely information (Bourne et al., 2002; Speckbacher et al., 2003; Brown, 2007, 9; Weil, 2007).

Lean Accounting

Borrowing liberally from English translations of Toyota's development of Lean operations, all the way from *The Machine That Changed the World* (Womack, Jones, and Roos, 1990) to *The Toyota Way* (Liker, 2004) and *The Toyota Way Field Book* (Liker and Meier, 2006), Lean accounting intends to adapt to accounting the basic principles of eliminating waste, reducing time and cost, and developing value streams.

Lean concepts were developed in the manufacturing industry, but even service industries now are adopting Lean techniques. For example, in an attempt to reinvigorate its operations, Starbucks recently introduced Lean techniques in its coffee shops (Jargon, 2009). An executive search firm (recruiting Lean executives) has begun using Lean concepts (Brandt, 2009), back offices are implementing Lean (Brewton, 2009), and even hospitals are trying it out (Does et al., 2009).

Connection to Value Stream Analysis (Cell Manufacturing Analogy)

The traditional accounting approach consists of gathering, by department, division, or segment, *direct costs*, which include all variable costs of production plus fixed costs benefitting a single unit, and allocating *common costs* (shared fixed costs) to all units that benefit from the common costs. In contrast, Lean accounting, like Lean

operations, focuses on establishing, for a *value stream* (a production flow for a particular product or family of products), a flow of data that rapidly produces high-quality information (Maskell and Baggaley, 2004, 9-10). For example, if processing accounting transactions individually, one by one, speeds up the flow of information to operating managers, that methodology is preferred even though batch processing of data may be a more cost-effective process. Most Lean accounting advice, though, applies to operations, not to activities of the accounting department itself.¹¹

Applying Lean accounting concepts to an operation where costs are aggregated by value streams, established for each product line or family of products, requires “dedicated” value stream resources. Each stream is designed to speed the flow of production and minimize arbitrary cost allocations. Dedicating resources to each stream results in some duplication of resources. Duplication of resources, of course, increases costs. Production, however, is speeded up and revenue is earned more rapidly.

Lean accounting recognizes the arbitrary nature of allocating common (shared) fixed costs and attempts to avoid this issue either by dedicating resources to individual value streams where allocations are limited to product family members or just not allocating common manufacturing overhead costs such as those for buildings, security, human resources departments, etc. Demonstrating the strong tendency to revert to accountants’ extensive traditional accounting education, however, two Lean accounting books (Maskell and Baggaley, 2004; Huntzinger, 2007, [Chapter 17](#), see especially p. 259) recommend allocating common fixed resources in order to produce a total cost per unit. A fully allocated cost per unit, though, is of dubious value to managers.

Advantages of Lean Accounting

Lean accounting proponents claim that through participating in *kaizen* events, an attempt to attain continuous improvement and referred to as a *kaizen blitz*^(SM)[12](#) when performed by a focused team over a short period of a few days, the exposed opportunities for improvements can be supported by accounting measures and reports. By understanding and reporting the results of Lean initiatives, Lean accounting supports improvements such as:

- Reduction, frequently dramatic, in work-in-progress (WIP) inventory[13](#);
- Elimination of non-value-added processes resulting in decreased total processing times;
- Increased company productivity;
- Reduced setup and changeover times;
- Increased on-time deliveries (Womack et al., 1990, 81; Liker, 2004, 3-6; Polischuk, 2009; Shipulski et al., 2009).

These advantages are the result of applying Lean concepts throughout an organization or a supply chain. Even with accounting support, realizing benefits promised by Lean initiatives is extremely difficult.

Disadvantages of Lean Accounting

General lack of success in copying another organization's strategy has been experienced, if not reported, by many firms. Attempting to reproduce improvement results on other than a short-term basis without supporting behavioral and cultural changes generally has not been successful. Failures of Lean implementations, and Lean accounting, have pointed to the following deficiencies:

- Top management does not actively and continuously support Lean initiatives;
- Accounting/finance people are not included in Lean training sessions (a not uncommon situation for all improvement initiatives);
- Information flows are not adapted to match new Lean flows (value streams);
- Publicizing local “successes” creates competition among various units of an organization;
- Appropriate performance metrics are not developed (Achanga et al., 2006; Stuart and Boyle, 2007; Pullin, 2009; Shook, 2009).

In addition to the new accounting developments discussed previously, some traditional accounting techniques such as standard costs and master budgets have remained powerfully in place.

Traditional Budgeting, Capital Budgets, and Control Mechanisms

While some organizations disparage the budgeting process (Anonymous, 2003; Hope and Fraser, 2003; Nolan, 2005; Weber and Linder, 2005), most companies still go through the annual angst of budget preparation with all the pomp, posturing, and political maneuvering of a public sporting event. Even if bad behaviors erupt, there are good reasons, such as detailed planning, company-wide coordination, and synchronization of effort, to undergo this process.

Regardless of the particular accounting methodology used to record and report transactions internally, most organizations, including governments, prepare budgets for various time periods, typically for a quarter or annual

period, but sometimes for three- or five-year periods. Budgets not only can be for differing time periods, but can be very specialized, such as a budget for a new product introduction or another individual project, an operating budget focusing on expected (or hoped for) operating income, a capital budget for asset acquisition, or a budget covering an entire organization, called a “master” budget.¹⁴

Master Budgets

Comprehensive (master) budgets should follow carefully laid strategic plans (although sometimes they proceed, or evolve into, the strategy). This process forces introspection and consideration of underlying assumptions and intended or possibly unintended consequences. In addition, budgets can project cash flow shortages in time to acquire bank-lending commitments at favorable times—before the cash is needed. For convenience, an annual master budget typically is broken down, somewhat arbitrarily, into monthly or quarterly subperiods and can require many months of back and forth communications between the finance department or budget committee and affected departments, business units, or segments (Bragg, 2007a, 30).

Financial planning includes preparing a projection of what the company hopes to accomplish for the next period. The typical budget process begins with projected sales in units and in currency, on a monthly basis, for a 12-month period.¹⁵ Based on expected sales and certain information concerning desired inventory levels, production in units, material acquisition (in units and currency), labor costs, variable overhead elements, and other production fixed overhead amounts to be incurred are estimated, both in terms of cash outflows and overhead applied, for each month of the period. At this point, cost of sales, including materials, labor, and applied overhead, is computed for

each month. Next, a schedule of general, selling, and administrative expenses, usually divided into variable and fixed components, is prepared.

Using the previous information, along with assumptions concerning collections from customers, payments to suppliers, asset acquisitions, and the timing of other cash inflows and outflows, a statement of changes in cash is prepared. Only then is sufficient information available for projected income statements and balance sheets. (See master budget relationships, in the diagrams, in Hilton, 2009, 350; Garrison et al., 2010, 375.)

Capital Budgets

One of the largest cash requirements involves acquisition of additional assets. In large, decentralized organizations, capital budget requests are prepared by *investment centers* that have responsibility for return on assets as well as profit and loss. These centers require additional resources in order to fulfill their stretch goals and do not (and probably cannot) predict the impact of their request on the entire organization. Therefore, executive committees usually schedule marathon sessions where managers come in and present their cases, using projected cash flows and net present values, for additional investment. The committee then must decide which proposals to fund¹⁶ based on logical analyses of short, compelling, and often competing presentations.

Due to the time required to achieve budget agreement on budgets of all sorts, management is frequently reluctant to revise budget numbers when original assumptions are invalidated. Thus, the data formulated during the fall for a calendar-year company may be used for the next 12 to 15 months.

Once the budget period begins, large organizations typically prepare periodic *flexible budgets*, which are based on actual results of the critical area of the organization on which a master budget is originated. For example, if a company has excess capacity, the first schedule in a master budget is sales units, followed by projected sales in the organization's currency, followed by production schedules, material schedules, etc. Therefore, once sales are known, expected costs associated with those sales can be prepared. If a company has more demand than it has resources to fulfill, however, the master budget must begin with a production schedule featuring the desired product mix. Then sales, in currency and other supporting schedules can be derived. Later, actual production and sales of various products, along with expected costs, can be combined in a flexible budget. Flexible budgets present more meaningful interlocking data once actual critical area performance is known.

Use of Budget Data

In addition to their planning role, budgets frequently are used in performance evaluation. Because a master budget contains budgets for each area of an organization, it is easy to engage in *performance to budget*, where actual results for each area are compared with budgeted predictions and favorable or unfavorable variances computed and reported. Flexible budgets, as reported by every cost and management accounting textbook produced in the last several decades, control some, but not all, of the damage of this type of performance report.

Advantages and Disadvantages of Traditional Budgets

Advantages of budget preparation include the following:

- Planning for future periods is required.
- Budgets facilitate communication throughout an organization.
- Goals (expectations) are set.
- All areas of the organization have input into the process.
- Upcoming creditor covenants can be met or renegotiated.
- Resource requirements can be established (Hilton, 2009, 348–349; Garrison et al., 2010, 369).

Disadvantages of the budgeting process and its uses are that it:

- Sets an upper bound on performance (diminished incentive to “outperform” the budget);
- Encourages padding of requests (gaming) in anticipation of forced reductions;
- Results in a lack of budget ownership (budget numbers frequently are dictated by upper-level management);
- Encourages competition between areas for resources;
- Can encourage dysfunctional behavior in order to meet budget amounts;
- Is cumbersome and too expensive (Cunningham and Fiume, 2003, 133–139; Hope and Fraser, 2003, [Chapter 1](#); Hilton, 2009, 375–376).

TOC Approach to Planning, Control, and Sensitivity Analysis

As amazing as it may sound, TOC makes the strong assertion that all the allocation gyrations of traditional and newer accounting methodologies are not necessary and generally serve to confuse and obfuscate rather than enlighten. In fact, implementing TOC (or any other management improvement initiative) without changing the internal accounting and reporting system will send mixed messages to the troops and eventually, by encouraging people to go back to old and out-of-date policies and assumptions upon which previous reporting is based, will undermine the new system.

Planning

At its most basic level, planning includes establishing strategy and then implementing the chosen strategy. Because this subject is treated in detail in later chapters in this Handbook ([Chapters 15](#), [18](#), [19](#), and [34](#)), treatment of strategy and tactics are deferred until later.

The typical starting point for planning in a Throughput environment is recognition of the organization's most binding constraint (Step 1 in TOC's *Five Focusing Step* [5FS]-process). If raw materials are in short supply, vendors may occupy this position. Most often, though, an organization's demand from its customers poses the most binding constraint, especially in times of recession in the economy as experienced in the last half of 2008 and in 2009. Because of company policies, however, it is not unusual also to find one or multiple internal constraints.

Finding the Best Product Mix

Accounting people typically think about capacity in terms of facility capacity, not the capacities of individual resources used to produce a company's products. If demand is greater

than any one of an organization's resource capacities, however, products must be prioritized.

The traditional approach is to prioritize products based on one of the following: (1) selling price, (2) gross profit, or (3) contribution (gross) margin. An activity-based accounting system prioritizes products based on activity-based gross profit for each product.

TA uses explicit recognition of an internal constraint when prioritizing products. One of the most familiar TOC formulas used to determine the best product mix when demand is greater than production capability (an internal constraint exists) is throughput per unit of constraint time. Accountants will have learned this concept under the name *contribution margin per unit of constraint*, which is recommended to determine product priorities when an organization faces a single constraint.¹⁷ This process most easily is illustrated with an example.

[Figure 13-1](#) is adapted from the original "P-Q" example developed and presented by Eli Goldratt in numerous workshops all over the world and in one of his books (Goldratt, 1990, [Chapter 12](#)). Rather than two products, [Fig. 13-1](#) has three products, but the basic idea of a stable environment with no significant uncertainties is the same. Given [Fig. 13-1](#) and some basic information as shown in Tables 13-1 and 13-2, a TOC-trained person can compute the optimal product mix (Product Z, then Product X, then Product Y) and expected operating income in a matter of minutes.

Three elements in [Fig. 13-1](#) have darker outlines because their output is required in more than one product. Resource 1, task 2, and Resource 2, task 3 produce a common component from Raw Material #3 that is used in both Product X and Product Y.

[Figure 13-1](#) shows a production view (combining both bills of materials [BOMs] and routings for items flowing through

the facility) of the organization’s operations where each of the four resources can perform different tasks. A typical accounting view would show the materials flowing through four stationary resources.

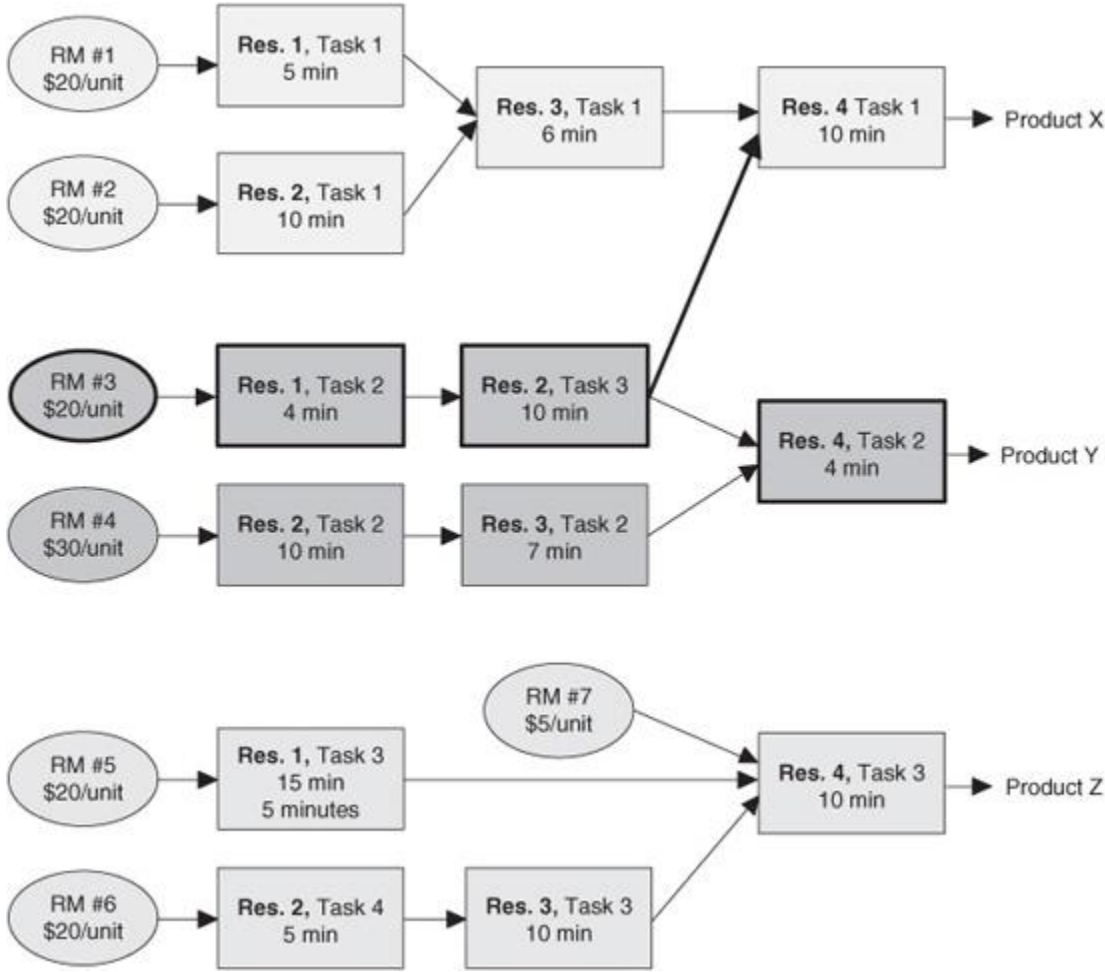


FIGURE 13-1 Product flows through resources for a simple company.

Within five minutes, most people familiar with TOC concepts recognize that Resource 2 does not have sufficient capacity to produce all units demanded and therefore would compute the Throughput (contribution margin) per minute required of Resource 2 as follows:

Product X: $\$300 - \$60 \text{ materials} - \$8 \text{ VMOH} - \$32 \text{ VSC} = \$200 / (20 \text{ min of Res. 2}) = \$10.00/\text{min}$

Product Y: $\$260 - \$50 \text{ materials} - \$5 \text{ VMOH} - \$22 \text{ VSC} = \$183 / (20 \text{ min of Res. 2}) = \$9.15/\text{min}$

Product Z: $\$195 - \$45 \text{ materials} - \$2 \text{ VMOH} - \$15 \text{ VSC} = \$133 / (5 \text{ min of Res. 2}) = \$26.60/\text{min}$

Product	Units Demanded per Week	Selling Price	Variable Mfg. Overhead (VMOH)	Variable Sales Commission (VSC)
X	90	\$300.00	\$8.00	\$32.00
Y	50	\$260.00	\$5.00	\$22.00
Z	80	\$195.00	\$2.00	\$15.00

TABLE 13-1 Demand, Selling Prices, and Variable Costs

Item	Availability or Cost per Week
Resources 1, 2, 3, and 4	Each resource available 2400 min per week (total of 9600 min per week)
Wages (shared among all products)	Cost of \$4,800 per week
Fixed (shared) manufacturing overhead	Cost of \$7,200 per week
Fixed (shared) general, administrative, and selling overhead	Cost of \$5,612 per week

TABLE 13-2 Additional Information

Method	Operating Income
Throughput (Z, X, Y)—constraint recognized and exploited	\$12,858
Traditional Gross Profit (Y, X, Z)—constraint not recognized	\$ 5,538
Traditional Contribution Margin (X, Y, Z)—constraint not recognized	\$ 5,878
Activity-Based Cost (X, Y, Z)—constraint not recognized	\$ 5,878

TABLE 13-3 Operating Profit Resulting from Various Accounting Priorities

Therefore, product priority would be Product Z, then Product X, then Product Y. Weekly income would be computed as \$12,858 (total Throughput—or contribution margin—of \$30,470 minus total fixed costs of \$17,612), using all 2400 minutes of Resource 2.¹⁸

Following reasonable assumptions,¹⁹ the traditional gross profit or gross margin approach would result in first priority going to Product Y, then X, then Z. (See a complete list of 13 assumptions, some of which we will not need for the examples in this chapter, in a spreadsheet, “Throughput_Examples”)²⁰ Similarly, ABC²¹ would result in gross profit priorities of Product X, then Y, then Z. [Table 13-3](#) compares the operating income (for simplicity, taxes are ignored) for four methods (Throughput, traditional, traditional contribution margin, and ABC).

Once the best product mix is determined, a formal master budget can be prepared.

Preparing a Throughput Budget

Throughput budgeting would follow the same general flow as that described in the section on traditional budgeting, but with conscious consideration of a possible internal constraint. The budget preparation process best proceeds when production provides the following data: (1) BOM for each product; (2) routing for each product; (3) prioritized expected sales of each product; (4) required inventory sizes; (5) available resource capacities; and (6) proposed acquisition of land, buildings, or equipment during the period.

With an internal constraint, the budget process would begin with the estimated production of the most profitable

product mix (in units), and consideration of constraint availability. Then estimated sales (in units and in total revenues), production costs, and all other elements of a traditional budget would be prepared. Following preparation of the cash budget, the income statement would be prepared in two formats: the direct costing format used by Throughput accounting^{22,23} and the traditional GAAP format (revenues minus cost of sales, subtotaled into gross profit, minus general, selling, and administrative expenses to find operating profit). (See a complete treatment of Throughput budgeting in Bragg, 2007b, [Chapter 5](#).)

The Throughput budget would be used for planning purposes only, and not for control as traditionally practiced.

Throughput Control

TOC maintains that three straightforward metrics—*Throughput, sales revenue minus all variable costs (manufacturing and general, selling, and administrative), Inventory (or Investment), the funds an organization has expended to be in position to produce, and Operating Expenses, the repetitive expenditures a company must incur each period in order keep the company operating*—are all the measures needed for day-to-day operating decisions (Goldratt and Cox, 1984). These three metrics have occupied space in every cost and management accounting textbook, under slightly different names, since at least the 1960s (Dopuch and Birnberg, 1969).

Different definitions of terms undoubtedly have caused much confusion. While it may be impossible, at this late date, to change TOC terminology, people trained in accounting call *Throughput* by the name *contribution margin*,²⁴ with the same definition—*revenue minus totally variable costs*. *Inventory* is a highly controllable subset of *Investment* in total assets. *Operating Expenses*, in

accounting terminology, would be *fixed costs*, including manufacturing fixed costs and general, selling, and administrative fixed costs.

Where to Focus Quality Improvements

The same example shown in [Fig. 13-1](#), along with identification of the internal constraint (Resource 2) can be used to focus quality improvements. Assume the company experiences a scrap problem at Resource 4 resulting in 4 percent of units (3.6 units) of Product X being scrapped, 7 percent (0.7 units²⁵) of Product Y, and 8 percent (6.4 units) of Product Z (see [Fig. 13-2](#)). The quality team can correct the problem on only one product at a time, the cost being approximately equal for each product. Which product should they work on first?

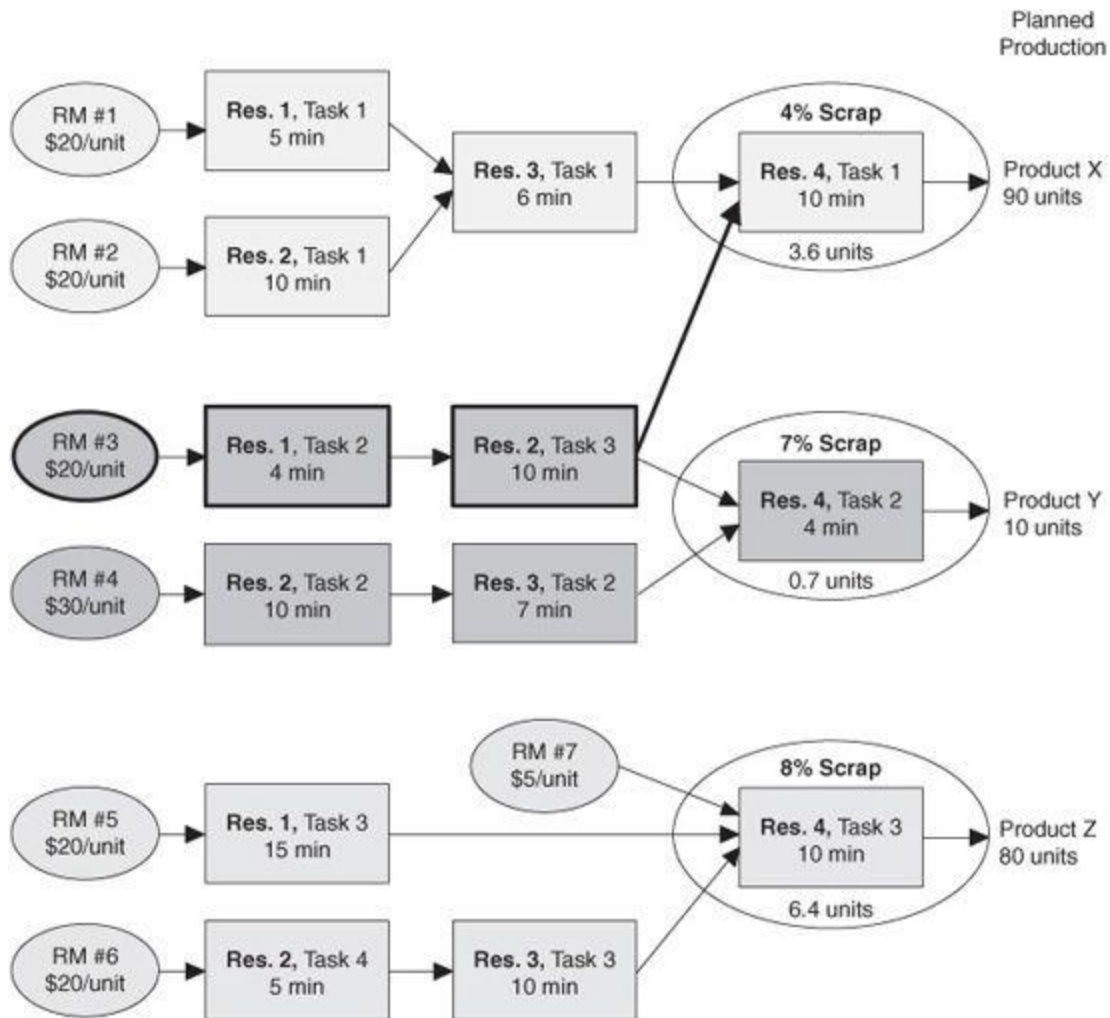


FIGURE 13-2 Scrap problem at Resource 4.

Provided the quality problem data shown in [Fig. 13-2](#), most people immediately select Product Z for first attention because (1) it has the highest percentage of scrap, (2) it has the most units being scrapped, or (3) it is the company's most profitable product. With traditional accounting, even if the cost of the time lost (45, 35, and 40 min, respectively, for Products X, Y, and Z), at \$0.50 per minute²⁶ is included in the analysis, along with the cost of materials (\$60, \$50, and \$45 for Products X, Y, and Z, respectively) and variable manufacturing overhead (\$8, \$5, and \$2 for Products X, Y, and Z, respectively), the priorities remain Product Z (with a

total cost of \$428.80), then Product X (total cost of \$325.80), and then Product Y (total cost of \$50.75).

Because there is sufficient time to replace the work lost on Resources 1, 3, and 4, however, the Throughput approach includes lost variable costs (materials and variable manufacturing overhead) plus the cost of lost time only on Resource 2. Units lost of Products Z and X will be replaced, resulting in fewer units of Product Y being produced and sold. First, we safely may ignore the option of fixing the quality problem on Product Y because only 1 unit is lost per week, leaving only Products X and Z for improvement consideration.

Since only whole units may be sold, if the quality problem on Product Z is corrected first, Product X's remaining quality problem will result in 4 fewer units of Product Y being produced and sold. Product X production losses will be replaced, but 80 min (4 unit \times 20 min) on the constraint (Resource 2) will have been lost. Thus, after Resource 2 has been used to produce 80 units of Product Z (still the best product, requiring 400 min) and 90 units of Product X (the next best product, requiring 1800 min), leaving only 120 min to produce and sell 6 units of Product Y, resulting in operating income of \$12,126.

However, if the quality problem on Product X is eliminated first, Product Z's remaining quality problem will result in 7 units having to be replaced, at 5 min per unit, meaning 35 min of Resource 2's time will be lost, leaving 2,365 min available. Once again, 80 units of Product X, requiring 400 min on Resource 2, and 90 units of Product Z, requiring 1800 min on Resource 2, leave 165 Resource 2 minutes available for the production and sale of Product Y. With 165 min available, 8 complete units of Product Y [(165 min)/(20 min per unit)] can be produced and sold. With this product mix, operating income will be \$12,492, \$366 higher than if the quality problem is fixed first on Product Z.

While looking only at the relevant resource minutes lost on each product, multiplied by Product Y's Throughput (contribution margin) per minute of Resource 2 production time, plus the variable costs involved will provide the same answer (fix the quality problem at Product X first, then go to Product Z's problem). This shortcut method (see the *Throughput Examples* spreadsheet, AR29: AZ48) risks failing to consider all variables. The total effect on operating income (shown in cells AQ50:BB69 in the *Throughput Example* spreadsheet) is by far the safer method and the one recommended by experts.

Adding a New Market Segment

The Throughput solution often has been compared with a linear programming solution of a one-constraint problem (Dopuch and Birnberg, 1969) and has been described as a step-wise linear programming analysis, dealing with one constraint (the worst one) at a time. Unfortunately, the Throughput solution, like a linear programming solution, is extremely sensitive to deviations from an *equilibrium* solution, one in which a best solution is found using current assumptions concerning resource availabilities, product demand, and so forth.

For example, suppose a salesperson returns from China with an order for 30 units a week for each of the three products, X, Y, or Z, or any combination thereof, with agreed selling prices equal to 80 percent of the U.S. prices. Should the company sell any of its products in China? Facing this decision, the company must be very careful not to make the easiest of mistakes: assuming the constraint will not shift to another resource.²⁷

After computing Throughput per minute of Resource 2 for each of the three potential China products, suppose the company decides to sell 30 units of Product Z in China

(called China Z in the spreadsheet) with a Throughput per minute of Resource 2 of \$18.80 ($\$156 - \$62 = \$94 \div 5 \text{ min on Resource 2}$), prior to filling orders for Products X and Y, and will not be interested in selling Product X (China X, with a Throughput of \$7/min) and Product Y (China Y, with a Throughput of \$6.55/min) in China. Following this strategy, however, will cause the company not to make a higher total profit (\$14,214), as it expects, but to make \$12,448—\$1,766 less than expected—and \$410 less than its *previous* best performance with no sales to China. The deterioration in operating income will occur due to Product Z's (and, therefore, China Z's) high usage of Resource 1, causing it to be in tighter supply and resulting in an interactive constraint with Resource 2. (See "Throughput_Example" spreadsheet, cells BD2: BS82.)

Controls should be in place to prevent actions that will reduce operating income. The following examples illustrate how traditional accounting can lead to nonoptimal decisions.

Purchasing Decisions

Even though materials are not often an organization's constraint, rapid expansion in 2007 and 2008 saw raw materials prices skyrocket. Of course, the recession in late 2008 and 2009 brought material costs back in line. When materials prices change, Throughput and Throughput per unit also change. Therefore, product priorities also may change. In a TOC world, any time any Throughput metric input changes, its impact on priorities must be computed.

Less obvious purchasing decisions involve opportunities to acquire materials from a lower-cost supplier or to outsource certain portions of the productive effort. Potential acquisition errors can occur based on both accepting and rejecting outsourcing proposals as well as on initial material purchases. Each of the following decisions should be

considered independently. That is, the starting point is the current most profitable combination of 80 units of Product Z, 90 units of Product X, and 10 units of Product Y.

Acquisition Decision Purchasing has found a new supplier who is willing to provide Raw Material #7, for \$2.50, saving the company \$200 a week. [Figure 13-3](#) illustrates this opportunity. If Purchasing primarily were evaluated based on cost savings, they would like to make the deal. After trying a sample of the new material, however, the production manager states that Resource 4, Task 3 will incur approximately 10 percent scrap. Since Resource 4 has plenty of idle time, Purchasing assures the production manager, they can easily make up the 80 min lost due to scrap. Further attempting to seal the deal, the purchasing person tells the production manager that since Resource 4's utilization will increase, efficiencies may increase, offsetting any scrap.

The production manager, having been trained in TOC concepts, states (not too patiently) that since the scrap occurs following processing on the constraint (Resource 2), each lost minute on Task 4 means that fewer units of other products can be produced and sold. Additional materials will be processed to make sure all demand for Product Z will be filled, so the lowest priority product, Product Y, will take the hit. Because Resource 2, Task 4 requires 5 min of processing time per unit, 40 min of Resource 2 time (8 units × 5 min) will be lost. Because Product Y requires 20 min per unit on Resource 2, two units of Product Y will be eliminated (at \$183 Throughput per unit). Therefore, this material cost "savings" of \$200 will cost the company \$366²⁸ in lost Throughput every week! (See the original "best" operating income versus the operating income if the proposed change were accepted in the *Throughput_ Examples* spreadsheet, cells BU1: CF40.)

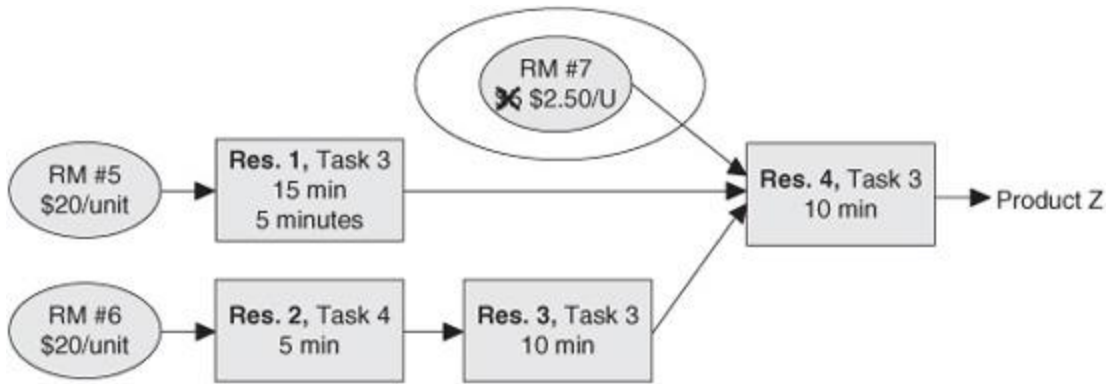


FIGURE 13-3 Proposed acquisition of a cheaper material.

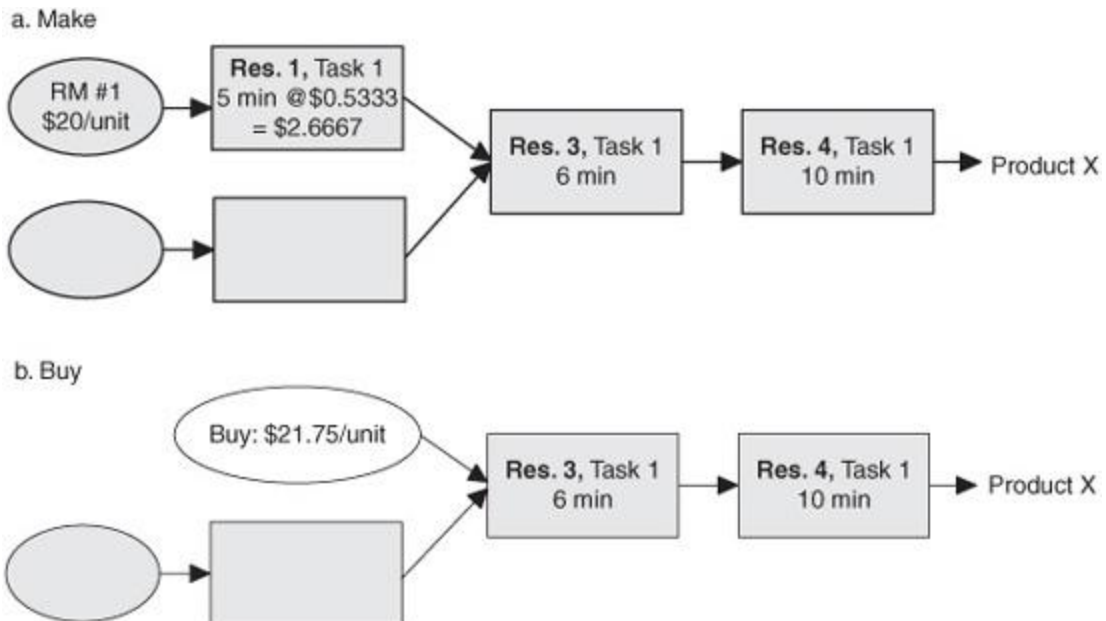


FIGURE 13-4 Product X, Resource 1, Task 1—make versus buy decision.

This “opportunity,” if accepted, would result in a decrease in operating profit of \$166 (\$366 - \$200) each week. Fortunately, the production manager rejected this “cost-saving” proposal.

Outsourcing Proposal #1 Suppose another person in purchasing has received an offer from a new supplier to

provide a component that will include Raw Material #1 and processing through Resource 1 for a cost of \$21.75. No variable overhead is incurred for this operation. Should the offer, shown in Fig. 13-4, parts a (make) and b (buy), be accepted? Traditional accounting would say, “Absolutely.” The unit cost through that point in production is \$22.67 (\$20.00 for the material and \$2.67 for 5 min of processing at a cost of \$0.5333 per min); resulting in a savings of \$82.80 each week (\$0.92 a unit times 90 units) or over \$4100 each year.²⁹

However, TA would respond, “No way!” Resource 1 is not a constraint and already has 50 min of unused time each week.³⁰ Accepting this outsourcing offer would result in incurring an additional cost of \$1.75 a unit (\$21.75 - \$20) for the 90 units needed—\$157.50 each week³¹; almost \$8000 a year. Meanwhile, Resource 1 would incur additional idle time of 450 min each week. In addition, the company would lose direct quality control and incur the risk of unavailability of the component when needed.

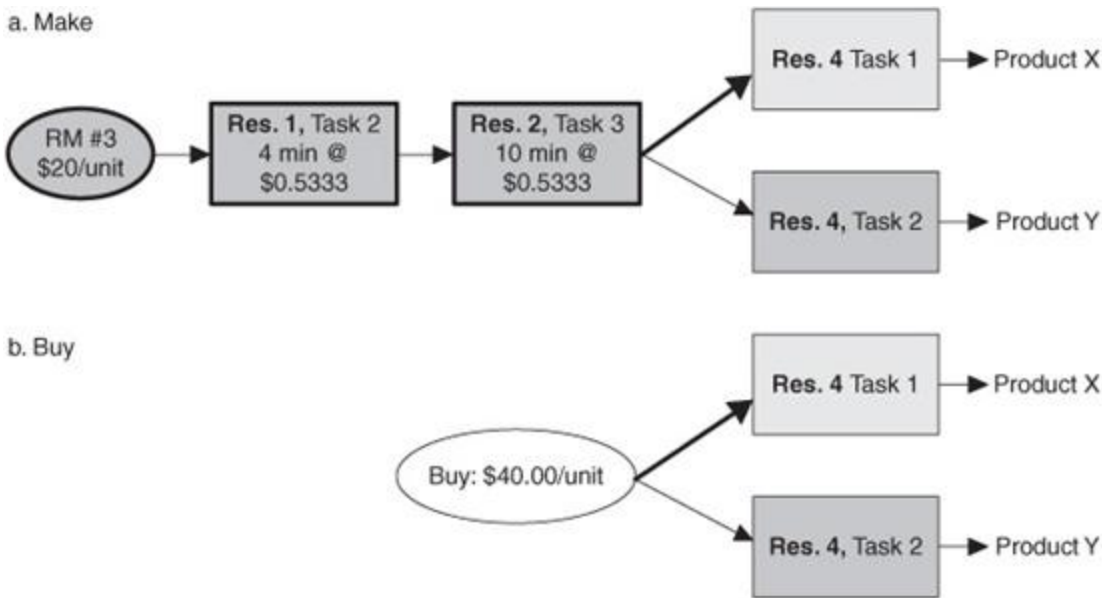


FIGURE 13-5 Products X and Y, Resource 1, Task 2 and Resource 2, Task 3—make versus buy

Of course, traditional accounting would respond that Resource 1 should be put on a 4-day workweek since they now have over 8 hours of idle time. Sometimes this makes sense, but not normally. Cutting the pay, effectively, of one worker does not inspire high worker morale and job commitment. In addition, Resource 1 is the most likely constraint candidate should Resource 2 capacity's be elevated.

Outsourcing materials and Resource 1 work, in the situation described, would *not* be a good decision.

Outsourcing Proposal #2 Purchasing also has an offer from a supplier to provide a component that would include Raw Material #3 and processing by Resource 1 and Resource 2 for a cost of \$40. Variable manufacturing overhead for the operations involved is \$2.50 a unit. [Figure 13-5a](#) shows the current arrangement, and [Fig. 13-5b](#) shows the “buy” proposal. Should the company accept the offer?

Traditional accounting would say, “No.” The cost to make is only \$29.97 (\$20 for the material, \$7.47 for the labor—14 min at \$0.5333,³² plus \$2.50 variable overhead). Buying the proposed component would increase the company's costs by over \$10 per unit and \$1000³³ each week.

Rather than comparing costs, however, a person who is aware of TA would look at the impact on the company's Throughput. Resource 2, Task 3, requires 10 min. Recall that the constraint in this system is Resource 2. Saving 10 min of Resource 2 time on Product X (90 units) and Product Y (10 units) amounts to 1000 extra minutes. With the extra time, additional units can be produced for the unfilled demand of Product Y. All unfilled demand for 40 units of Product Y can be produced and sold, adding an additional \$183 per unit for a total of \$7,320 additional Throughput, and adding \$4520 to the bottom line (\$7320 - \$2800—the added cost of outsourcing 140 units³⁴ at \$20 incremental cost). Compared

to the example company's previous best performance of \$12,858, this change represents about a 35 percent increase.³⁵ Even if the company incurs additional overhead to track the supplier's quality and dependability, the outsourcing (buy) offer should be enthusiastically accepted. In addition, everyone in the company should be made aware of the fact that marketing is now the organization's constraint and management should formulate plans to increase the capacity of both Resource 2 and Resource 1 when product demand increases.

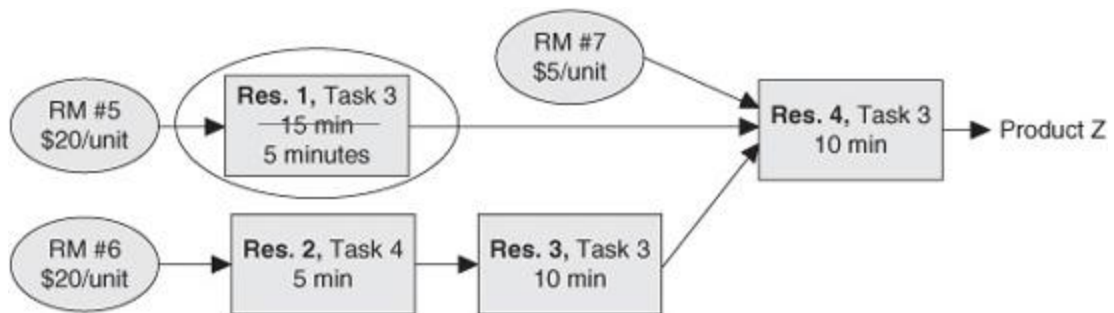


FIGURE 13-6 First engineer's engineering change proposal.

Engineering Change Proposals (ECPs)

Engineers have been studying production operations and two new engineers have submitted *engineering change proposals* (ECPs).

First Engineer's ECP A young engineer has read about a new process that can reduce the time on Resource 1, Task 3, from 15 to 5 min (see [Fig. 13-6](#)). New tooling costing \$5000 would have to be acquired, however. Should the proposal be accepted?

Traditional accounting typically would value this opportunity as favorable since saving 10 min on 80 units would produce a savings of 800 min. At the applied labor

rate of \$0.5333 (or the actual rate of \$0.50), cost savings would total \$426.64 or a minimum of \$400 a week. Thus, payback would occur in $\$5000/\$400 = 12.5$ weeks, at the longest. This is a very quick *return on investment* (ROI).

Of course, traditional accounting information provides support for an incorrect decision. Resource 1 is not the constraint and the “cost savings” of \$400 or more per week would never occur. Resource 1 would just have more idle time and the company would be out \$5000 for the tooling. TA correctly and immediately would reject this proposal.³⁶ Because Resource 1 might someday become a constraint (it has the highest loading after Resource 2, the current constraint), this proposal might be kept on file for action later, but not now.

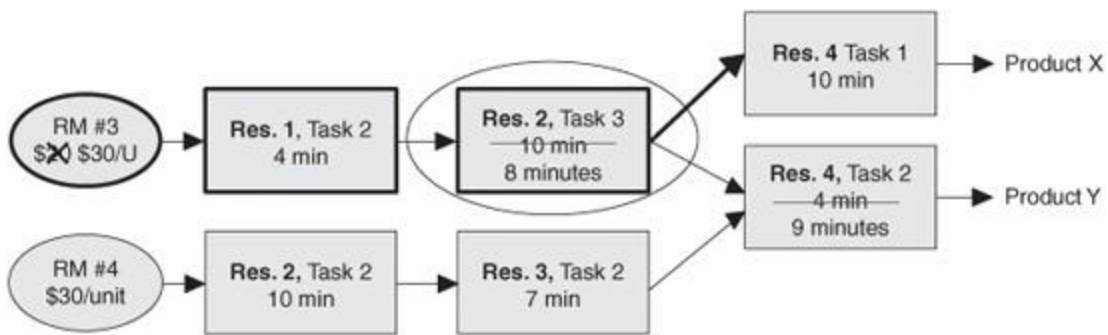


FIGURE 13-7 Second engineer’s engineering change proposal.

Second Engineer’s ECP Another engineer has submitted an ECP affecting Product X and Product Y. [Figure 13-7](#) shows three changes: (1) an increase in materials cost from \$20 per unit to \$30, and (2) a 2-min decrease on Resource 2, Task 3, and (3) an increase from 4 min to 9 min for final processing of Product Y by Resource 4, Task 2. Oh, and the change will require an additional \$8000 investment.

Of course, the accounting department is shocked by the \$10 increase in material cost and the 3-min increase in net

processing time for Product Y (from 35 to 38 min) that is partially offset by the 2-min decrease in Product X processing time. Considering the additional investment required, accounting might even suggest that this engineer should work for a competitor.

By now, you know that since Resource 2 is the company's current constraint, this change should be evaluated further using TA concepts. Saving only 2 min on Resource 2, Task 3, for 100 units (90 for Product X and 10 for Product Y) means an additional 200 min of availability but costs an additional \$1000 in raw material cost (\$10 per unit \times 100 units for the original quantities) in addition to the \$8000 investment. With this additional time, however, more units of Product Y can be produced and sold. How many additional units are possible? Not 10 (200/20), but 11 (200/18). With an original Throughput per unit of \$183, 11 additional sales of Product Y will bring in \$2013 each week. This amount, minus the \$1000 in additional cost for Raw Material #3, means an additional \$1013 in operating profit. Resource 4 has sufficient available time not only to use 9 min of processing time on the 11 additional units, but for the entire market demand of 50 units. The payback period for this investment would be \$8000/\$1013, or less than 7.9 weeks. Assuming the cost of the investment can be amortized over 52 weeks, operating income will increase by \$3986 each week.³⁷ This is a great investment, but it would have been turned down using traditional accounting metrics!

The Problem of Identifying Decision-Relevant Costs and How to Avoid a Disaster

Throughput (contribution margin), Inventory (investment), and Operating Expense (fixed costs) changes are always relevant. However, it is extremely difficult to accurately select the relevant costs and revenues (including those associated with lost opportunities) of many, if not most,

management decisions. For example, when multiple changes are occurring to more than one element of a process (the second engineering ECP is an example), keeping everything straight for a correct analysis can be difficult. The advice provided throughout the book, *Supply Chain Management at Warp Speed* (Schrage et al., 2009), is to consider *any* change (product mix, investment, make versus buy, special orders, rationalization of product lines, etc.) in terms of its impact on *total* amounts of Throughput, Inventory, and Operating Expenses (contribution margin, investment, and fixed costs, in accounting terminology). This same advice, couched in terms of the dangers of allocating fixed costs, is included in virtually every cost and management accounting textbook (see, for example, Hilton, 2009, 600-601, 612; Garrison et al., 2010, 588-589), and should be followed without exception to avoid costly errors.³⁸

Inventory Changes and GAAP Accounting

Basic goals of TOC are for Throughput to increase, Inventory to decrease, and Operating Expense to decrease. Throughput increases and expense decreases will be reflected favorably on external reports that conform to GAAP. Inventory reductions, however, will be reflected unfavorably on GAAP statements by reducing both assets and operating income. Therefore, inventory reductions should be handled with special care.

Because some accounting and other people have trouble understanding exactly how reducing inventory results in decreased income, I have developed several examples over the years to validate this result.

For example, assume a company that has no beginning inventories of WIP or finished goods, produces 20,000 units and sells 15,000 units for \$20 each. There is no ending WIP

inventory. Budgeted costs (as traditionally prepared) include the following:

Cost Item	Details	Total	Per Unit
Direct materials	40,000 units @ \$1	\$40,000	\$2.00*
Direct labor	2,500 hours @ \$10	25,000	1.25*
Var. mfg. OH	4,000 machine hours @ \$5	20,000	1.00*
Fixed mfg. OH	4,000 machine hours @ \$20	80,000	4.00*
Total product cost per unit			<u>\$8.25</u>
Var. sell. and admin.		30,000	2.00**
Fixed sell. and admin.		<u>75,000</u>	5.00**
Total costs incurred		<u>\$270,000</u>	

*Based on 20,000 units produced.

**Based on 15,000 units sold.

A traditional (absorption costing) income statement and a Throughput (variable or direct costing) income statement (both assuming costs are the same as those projected) are shown in [Fig. 13-8](#).

As shown in [Fig. 13-8](#), the traditional income statement shows net operating income of \$71,250, while the Throughput income statement produces net operating income of only \$51,250. The difference of \$20,000 (\$72,250 - \$51,250) can be reconciled solely by the change in inventory fixed costs. That is, the increase of 5000 units in finished goods times the fixed manufacturing cost of \$4.00 per unit equals the \$20,000 increase in traditional (GAAP) income over the Throughput income of \$51,250.

Traditional Income Statement			Throughput Income Statement		
Revenues (15,000 units @ \$20)		<u>\$300,000</u>	Revenues (15,000 units @ \$20)		<u>\$300,000</u>
Cost of Goods Sold			Variable costs		
Beginning finished goods	- 0 -		Beginning finished goods	- 0 -	
Direct material used	\$40,000		Direct materials	\$ 40,000	
Direct labor (all variable)	25,000		Direct labor (all variable)	25,000	
Variable mfg. overhead	20,000		Variable mfg. overhead	<u>20,000</u>	
Fixed mfg. overhead	<u>80,000</u>		Var. cost of goods mfgd.	\$ 85,000	
Total cost of goods mfgd.	\$165,000		Ending finished goods**	<u>21,250</u>	\$ 63,750
Ending finished goods*	<u>41,250</u>	<u>123,750</u>	Variable sell. and admin.		<u>30,000</u>
Gross margin		\$ 176,250	Total variable costs		<u>93,750</u>
Selling and Administrative Expense			Throughput (Contribution Margin)		
Variable	\$30,000				<u>\$ 206,250</u>
Fixed	<u>75,000</u>	<u>105,000</u>	Fixed costs		
Net operating income		<u>\$ 71,250</u>	Manufacturing		\$ 80,000
			Labor? (if it is fixed, the \$25,000 would be here)		
			Selling and administrative		<u>75,000</u>
			Total fixed costs		<u>155,000</u>
			Net operating income		<u>\$ 51,250</u>

* 5,000 units @ \$8.25 (variable and fixed manufacturing costs)
** 5,000 units @ \$4.25 (only variable manufacturing costs)

FIGURE 13-8 Traditional and Throughput income statements.

A more detailed example involving materials and WIP changes, as well as changes in finished goods, can be found in a spreadsheet entitled "InventoryReductionExample" located at www.mhprofessional.com/TOCHandbook. In this example, everything other than WIP inventory and *finished goods* (FG) inventory are held constant over a 4-year period. Year 1 sets up a baseline performance, the resulting income, and balance sheet. With no changes in inventories (Year 1), both methods (traditional and Throughput) produce the same net income before income taxes. WIP and FG inventories then are reduced by 50 percent in Year 2. Three spreadsheets are included in this file: "GAAP Accounting," "Throughput (Variable) Accounting," and "Reconciliation of GAAP & T. Inc."

In the example, a normal year (Year 1) income is \$200,000 and return on sales is 10 percent. The 50 percent WIP and FG inventory reductions occur in the second year (Year 2) and GAAP income drops to \$66,667 and return on sales drops to 3.33 percent. (See the “GAAP Accounting” spreadsheet, cells A48:Y80.) The third year of the example, GAAP return on sales recovers to 6.67 percent (cells A83:Y113) with income of \$133,333, but is not back to the full 10 percent (income of \$200,000) until Year 4 (A116:Y147). The reduced income occurs because, with traditional (GAAP) accounting, WIP contains a portion of fixed manufacturing costs, depending on the percent complete, and FG contains its fair portion (100%) of full fixed manufacturing costs. In the environment established in this example, the only way to reduce inventory is to cease entry of raw materials into the system.³⁹ The decrease in production activity required to lower inventories means that all fixed costs of the current period, plus the fixed costs in units in beginning FG and WIP, are charged to cost of sales in the inventory reduction year.

The TA approach, shown on the second spreadsheet of the “InventoryReductionExample” shows that income and return on sales for the entire 4-year period remain constant at \$200,000 and 10 percent, respectively. The third spreadsheet in the example file reconciles GAAP and Throughput income for the year of the inventory reduction and suggests general journal entries to adjust from internal Throughput reporting to external GAAP statements. [Table 13-4](#) shows the reconciliation followed by the general journal entries.

Because the inventory reduction is permanent, other things remaining equal, reported GAAP income would remain \$200,000 less than that reported under TA.⁴⁰ Given that this inventory reduction permits the opportunity to increase future earnings (lower WIP means faster processing

that permits additional production with no increase in fixed costs), the potential “sacrifice” in reported earnings is necessary and must be undertaken. Careful planning and communication with relevant stakeholders, especially employees, creditors, and owners, can minimize potential negative effects.

[Table 13-4](#) illustrates how the difference between GAAP income of \$66,667 and Throughput income of \$200,000 in Year 2 (a difference of negative \$133,333) may be explained totally by the change (reduction) in fixed costs in beginning and ending WIP of \$53,333 plus the change (reduction) in fixed costs in beginning and ending FG of \$80,000.

Following [Table 13-4](#) are all the year-end adjusting general journal entries to convert all income and balance sheet accounts from Throughput to GAAP. This example clearly indicates that keeping accounting records using Throughput concepts during a period quite easily can be converted to GAAP accounts at the end of a period.

Traditional Costing (GAAP) Income, Year of Inventory Reduction:	\$ 66,667
Throughput (Direct or Variable) Costing Income, Year of Inventory Reduction	<u>200,000</u>
Total difference to be explained	\$(133,333)

Difference explained by change in fixed costs in inventories:

Beginning WIP	Variable Costs	Fixed Costs	Totals
Materials	\$240,000		\$240,000
Labor	4,000	\$16,000	20,000
Manufacturing overhead	<u>\$16,000</u>	<u>\$144,000</u>	<u>\$160,000</u>
	\$260,000	\$160,000	<u>\$420,000</u>

Ending WIP	Variable Costs	Fixed Costs	Totals
Materials	\$120,000		\$120,000
Labor	\$2,000	\$10,667	\$12,667
Manufacturing overhead	<u>\$8,000</u>	<u>\$96,000</u>	<u>\$104,000</u>
	\$130,000	\$106,667	<u>\$236,667</u>

Change in fixed costs in WIP Inventory (\$160,000 – \$106,667)

\$53,333

Beginning FG	Variable Costs	Fixed Costs	Totals
Materials	\$180,000		\$180,000
Labor	\$6,000	\$24,000	\$30,000
Manufacturing overhead	<u>\$24,000</u>	<u>\$216,000</u>	<u>\$240,000</u>
	\$210,000	\$240,000	<u>\$450,000</u>
Ending FG	Variable Costs	Fixed Costs	Totals
Materials	\$90,000		\$90,000
Labor	\$3,000	\$16,000	\$19,000
Manufacturing overhead	<u>\$12,000</u>	<u>\$144,000</u>	<u>\$156,000</u>
	\$105,000	\$160,000	<u>\$265,000</u>

Change in fixed costs in FG Inventory (\$240,000 – \$160,000)

80,000

Total Change in Beginning and Ending Inventory Fixed Costs
(income differences fully explained)

\$133,333

TABLE 13-4 Reconciliation of Traditional (GAAP) Costing Operating Income and Throughput (Variable) Costing Operating Income for Year of Inventory Reduction

End of period adjusting entries to convert from Throughput to GAAP accounting:

Work in process inventory	106,667	
Finished goods inventory	160,000	
Fixed selling, general, and administrative expenses and production fixed costs		266,667
<i>to adjust WIP and FG balances to their GAAP, fully absorbed, amounts</i>		
Selling, general, and administrative expense	264,000	
Fixed selling, general, administrative expenses and production fixed costs		264,000
<i>to move fixed S, G, and A expense to its GAAP period account</i>		

Cost of sales	533,333	
Fixed Selling, general, and administrative expenses and production fixed costs		533,333
<i>to remove remaining manufacturing fixed costs from the "periodic" expense account to cost of sales</i>		
Cost of sales	400,000	
Deferred fixed manufacturing costs		400,000
<i>to adjust cost of sales and close the deferred fixed manufacturing costs account</i>		

Value Metric Used to Track Performance

To provide timely feedback to managers and operations personnel, TOC has some unique metrics that reveal both what should be done and what should not be done. These metrics support standard TOC policies and are designed to encourage appropriate behavior. Many use a TOC concept called *value days* that aggregates the value of amounts invested or delayed using the following formula:

$$V_n = V_{n-1} + \Sigma \text{Value}$$

where V_n = the value for a current time period (e.g., day or week)

V_{n-1} = the value for the previous time period

ΣValue = the total net value (\$ in the U.S.) invested or realized in the current period (e.g., days, weeks, months)

This formula basically says that every amount of currency invested for a day results in a lost opportunity to use that amount for some other purpose, and every day the amount is not recovered repeats the lost opportunity. Therefore, the value is not fully recovered until sufficient amounts have been received to cover the entire deficit. The basic idea is to age investments analogous to the aging of accounts receivable.

Inventory Value Days For example, if \$100 is invested in inventory on Day 1 and the inventory is not sold until Day 10, inventory value days would equal $\$100 \times 10$ days or \$1000. In this way, slow moving inventory is highlighted and the information provided encourages the quick sale of older inventory. It also provides a way to project demand saturation so that additional units are not manufactured or acquired. With thousands of products, the process becomes more complex, but can be accomplished either with enterprise software or with spreadsheets. A one-product example that includes acquisitions and sales over a period of 33 days (“VALUE-FORMULAandEXAMPLES.xlsx,” second spreadsheet, “Inventory Example”) may be found at the following Web site: www.mhprofessional.com/TOCHandbook. [Figure 13-9](#), based on the spreadsheet example, shows the difference between the absolute amounts invested, as traditionally recorded (dark color) and inventory value days (lighter color). The rapid growth in inventory value days beginning on Day 23 signals management that inventory is building too rapidly and gives advance warning to cut acquisition of this item, which happened on Day 29. While the increase also is signaled by the traditional inventory values, upon close inspection, it is not nearly as noticeable and dramatic.

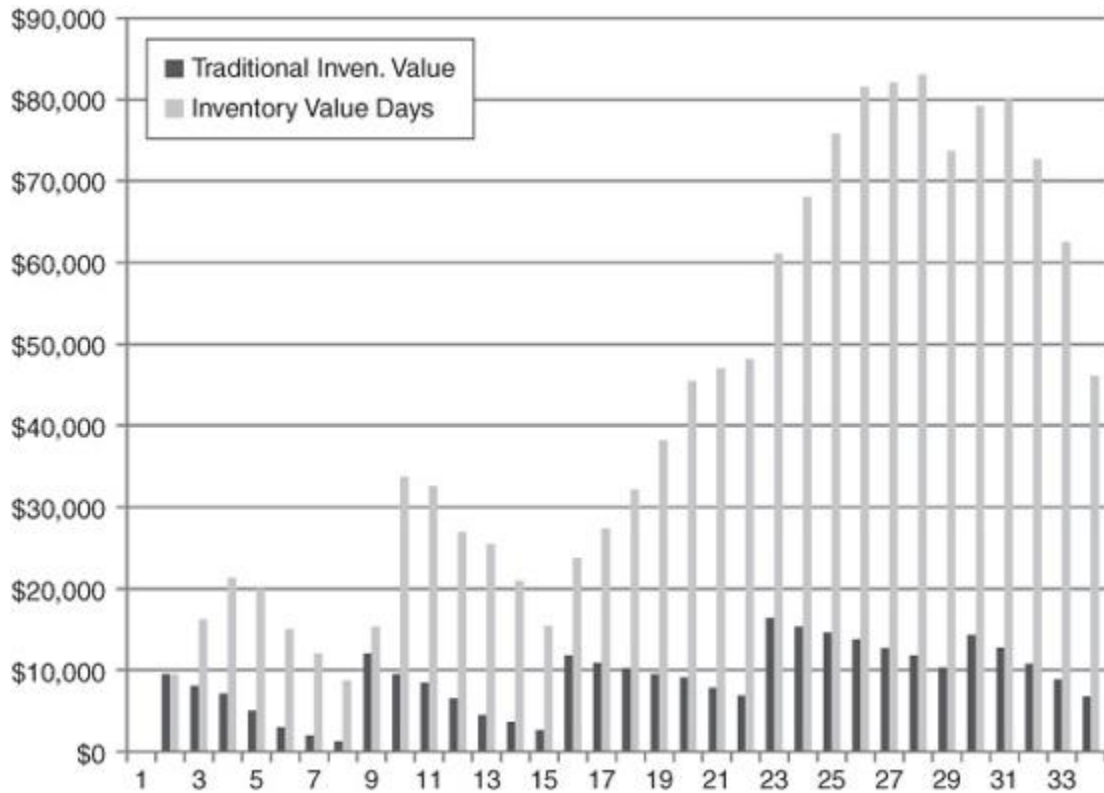


FIGURE 13-9 Traditional inventory value versus inventory value days.

Throughput Value Days The same file (“VALUE-FORMULAandEXAMPLES.xlsx”) has a third spreadsheet, “Throughput Example,” that illustrates how the same formula may be used to track orders that are delayed (missing from the critical area of a buffer for quality or other reasons or returned by the customer due to a quality problem).

Transfer Pricing

Organizations regularly go through restructuring programs where they go from a centralized organizational structure, where divisions are cost centers and major decisions such as product selection, pricing, and investments are made at

headquarters, to a decentralized structure, where divisions are profit centers or investment centers. As their names imply, a *cost center* is responsible only for controlling costs; a *profit center* has responsibility for generating revenues and controlling costs, and *investment centers* are encouraged to behave as entrepreneurs and are responsible for making investments, generating revenues, and controlling costs. Organizations also regularly go through the process in a reverse direction: from decentralized profit and investment centers to centrally controlled cost centers.

Top management would like to control all decisions (a centralized structure), but they realize that they are too far from the action to act sufficiently to quickly compete in a meaningful way. However, when organizations are decentralized, incorrigible transfer-pricing issues arise if any products are transferred from one division to another, which they usually are.

Additional complications arise because division managers usually are in competition with each other. The selling division wants a higher price; the purchasing division wants a lower price. Transfers between divisions in different countries bring complex tax issues into the mix.

Addressing transfer pricing in detail is beyond the scope of this chapter. However, if you are interested in a nice treatment of transfer pricing in a TOC context, without international transfers, I recommend, "The 'Transfer Prices' Problem," in *Approximately Right, Not Precisely Wrong* (Eden and Ronen, 2007, 241-258).

Other TOC Metrics

An ideal TOC operation has minimal fire fighting and expediting, no chaos or brutal overtime, and sufficient orders with reliable promise dates entering the system. In order to accomplish this desirable environment, it is critical

that sufficient raw materials are on hand,⁴¹ minimal (but sufficient) WIP and FG inventories are held, orders are delivered on time, appropriate buffers are established, buffer penetration (consumption) is tracked, and performance is continuously improving. In addition to the previous metrics, some general measures that organizations have found useful include the following:

- On-time deliveries—1 minus the ratio of late orders, weighted by days late, divided by total orders.
- Throughput per employee—Throughput (revenue minus variable costs) divided by number of employees.
- Inventory turns—Variable cost of sales divided by average Inventory held during the period.⁴²
- Throughput per unit of Operating Expense—Throughput divided by Operating (fixed) Expenses.

Of course, specialized industries have developed many other TOC metrics to provide feedback and control information in real time, or close to real time.

Sensitivity Analysis

The same metrics used for Throughput control can be used to perform “What if?” types of calculations. Sometimes decisions must be made quickly without time for submission to another person or department for analysis. In order for operating people to be able to analyze a situation quickly, they need a simple-to-understand method. For example, the same “value days” formula illustrated earlier for Inventory and Throughput control can be used to compare various possible investments. Four examples of \$60,000 investment opportunities, each with different cash flows, are shown in

the “Investment Examples” spreadsheet of the file “VALUE-FORMULAandEXAMPLES” (located online at: www.mhprofessional.com/TOCHandbook).

A “flush” point, where all value days’ cash investments have been recovered, follows Goldratt’s (1997, 246) development and is contrasted with payback period.

For comparison purposes, *net present value* (NPV) calculations are shown alongside the value-days calculations. For short-term (less than one year) decisions, NPV does not change significantly when using discount rates ranging between 10 and 20 percent. While the decisions indicated by NPV match those of the value-days calculations for three of the investment opportunities, although with much less clear discrimination, in one example, NPV indicates an acceptable investment and the value-days analysis shows it is unacceptable. The rationale behind using value days rather than NPV is that investment amounts are constrained by availability, not interest rates.⁴³

Throughput Accounting Approach to Performance Evaluation

TOC is very much a team sport. Therefore, evaluation should be on a team basis. Individual evaluation should be in the form of mentor feedback for improvement, as indicated by a request from the individual involved or a supervisor. In the absence of a clear need, feedback provided to individuals should not be used to evaluate performance. If a team is downgraded due to performance of one or more individuals, the team is expected to correct the problem. One company in Austin, Texas, hires all employees on a temporary basis for three months. At the end of the trial period, the team meets and decides whether a permanent position is offered.

Possible Explanations for the Lack of TOC Literature in Accounting and Finance

Three major reasons account for a dearth of TOC literature in accounting and finance. First, accounting students generally are not well trained in internal reporting and operations management. I believe that this may be because the *Certified Public Accounting* (CPA) professional examination has contained only about 10 percent of internal reporting material for a number of years. A new *Content Specification Outline* (CSO)⁴⁴ for the CPA exam, effective January 1, 2011, contains updated requirements that include increased emphasis on management accounting and operations, including TOC.⁴⁵ PhD students typically are not exposed to TOC, while MBA students at least are required to read *The Goal* (Goldratt and Cox, 1984).

Second, there is no incentive for accounting and finance professors to spend time becoming proficient in TOC concepts because they are primarily evaluated and promoted based on their research publications in the recognized top quality (mostly theoretical) journals where, at least in the United States, “applied” articles generally are not welcomed.⁴⁶ Bill Ferrar, recipient of the Lifetime Contribution to Management Accounting Award, has called for the teaching of TOC, but he suggests it could only occur by team teaching with a manufacturing engineering professor (Ferrara, 2007, 172).

Third, there is little or no demand from business school constituencies that TOC concepts be taught to students. Auditing firms frequently make accounting departments aware of their desire to hire students schooled in certain topics, such as XBRL (eXtensible Business Reporting Language) and IFRS (International Financial Reporting Standards), but as yet there is little or no push from industry

for accountants who have been trained in either TA or TOC. This should not be surprising since even firms that have adopted TOC typically do not include accounting and finance people in their training classes until improvement initiatives become accounting and finance targets for cost reduction.

Future TOC Accounting/Finance Research Needs

There have been relatively few publications by accounting or finance writers relating their research and conclusions on the subject of TOC. The field is entirely open.

Case Studies and Simulations

There is a desperate need for practitioners of TOC to partner with accounting or finance academics and publish case studies of their experiences, both good and bad, from a finance or accounting perspective, along with analysis of major factors that contributed to the success or failure of the initiative. It also would be most interesting to read a case of an accounting or finance area applying TOC concepts to their own operations. The quality and efficiency of accounting reporting systems could also be examined.

Information and Decision Making

Accountants have a broad perspective of an organization. To provide increased value to an organization, accountants need to establish internal information systems that aid decision makers. There is a need for research on the decision-making process, the behavioral aspects of decisions, single versus multiple decision makers and the

quality of decisions, and supply chain information and accounting.

Decision-Making Processes

While accounting and finance personnel usually do not make operating decisions, they do guard the treasury and must approve acquisitions. Therefore, they must understand legitimate investment needs.

It would be enlightening to see a rigorous study of short-term decisions, generally defined as decisions where the impact is experienced in one year or less and decisions must be made quickly, and long-term decisions where the decision time frame is longer and the impact is felt perhaps years later. What about the assumption that short-term decisions can affect (or become) the long-term? Who typically receives credit for long-term decisions where the cost is incurred much earlier?

What really is needed is a TOC conceptual framework for management accounting similar to Concepts Statements⁴⁷ provided by the Financial Accounting Standards Board (FASB) for financial accounting. That is, a TOC conceptual framework establishing desired information concepts would encourage an entire, internally consistent, reporting system including information objectives, decision support criteria, and periodic income statement reporting in a format that facilitates quick decisions by line managers as well as executive managers. A TOC conceptual framework would guide possible courses of action under various circumstances and promote the inclusion of the impact of decisions on external financial statements as well as on cash flows. Such a conceptual framework would guide the development of policies, procedures, and metrics to support a TOC environment.

Behavioral Aspects of Decisions

Several aspects of the behavior of decision makers can be influenced by the motivation and reward structure established by an organization's performance evaluation system. A study of the unintended effects of performance evaluation and how to structure a system that does not encourage dysfunctional behavior⁴⁸ such as that reported by Austin (1996) in *Measuring and Managing Performance in Organization* would be a significant contribution.

Because TOC requires teamwork, should evaluation be made only on the team's performance? How can a team address conflicts in local measures early in a change process? Should a team evaluate individual team members? Are team decisions superior to ones made by individuals? Always? Are certain decisions best made by an individual? How should decisions be evaluated?

Supply Chain Accounting

With all the changes being made in the structure and behavior of supply chains, it would be most interesting to see an accounting/finance solution to the problem of risk- and profit-sharing among all supply chain partners. Other issues, such as required quality, speed, and who benefits from reduced costs or other improvements, also need to be addressed.

Part of the charge for management accounting includes providing a comprehensive internal information system. During the recent global recession, supply chain partners several steps away from the final customer were shocked when markets suddenly dried up with no warning. Sometimes a supply chain member did not even know the final use of their component (Dvorak, 2009). The communication and presentation of information, both frequency and mode, merits further study.

Summary and Introduction of Remaining Chapters in This Section

Chapter Summary

To explain our current accounting and finance environment, the first part of [Chapter 13](#) describes a short history of cost accounting and the massive changes in the business environment during the 20th century. Management accounting's response, while lagging changes in business, includes the development of direct (variable or contribution margin) income statements to more closely tie income to sales, activity-based costing to more "accurately" trace all costs to cost objects (products), balanced scorecards to stress the importance of nonfinancial metrics, Lean accounting to match the value stream flows in manufacturing, and updated budgeting concepts that permeate all of these methodologies. Both the advantages and disadvantages of these approaches are reviewed.

The remainder of the chapter covers TOC concepts of planning, control, and sensitivity analysis. [Figure 13-1](#) establishes a simple example that is used to demonstrate both planning and control concepts. The negative impact of inventory reduction on GAAP accounting income is demonstrated and compared with TA results. Throughput *value days* is discussed and applied to inventory control, delayed Throughput, and potential investments. Additional TOC metrics are mentioned. Three files containing multiple spreadsheets, available online, provide supporting data for examples used in the chapter.

Finally, the chapter addresses possible reasons for the lack of TOC literature in accounting and finance and issues a call for further research.

Other Chapters Dealing with Performance Measures

In [Chapter 14](#), Debra Smith and Jeffrey Herman further describe desirable logistic measurements and demonstrate a framework to pull required information from an operation. They use TOC Thought Processes (TP) tools to defuse conflicts and deal with potential negative outcomes before they occur. A nice case study shows the application of these elements.

In [Chapter 15](#), Alan Barnard establishes a framework for designing and implementing a continuous improvement process, along with an auditing process to focus improvements where they are most needed. Alan describes the use of Strategy and Tactics trees both in implementing an improvement and in auditing the progress of the implementation.

[Chapter 16](#) provides a historical perspective on the need for a holistic approach to implementing TOC concepts. Two well-known TOC experts, Dr. Alan Barnard and Mr. Ray Immelman, present implementation case studies, one involving a public sector company and the other a private company.

References

- Achanga, P., Shehab, E., Roy, R., and Nelder, G. 2006. "Critical success factors for lean implementation within SMEs," *Journal of Manufacturing Technology Management* 17(4):460.
- Angel, R. and Rampersad, H. 2005. "Do scorecards ADD UP?" *CA Magazine* 138(4)(May):30.

- Anonymous. 2003. "How Alhsell discarded its budgeting process," *IOMA's Report on Financial Analysis, Planning & Reporting* 03(8)(Aug):11.
- Anonymous. 2008. "Linking strategy to operations," *Journal of Accountancy* 206(4)(Oct):80.
- Antonelli, V., Boyns, T., and Cerbioni, F. 2006. "Multiple origins of accounting—An early Italian example of the development of accounting for managerial purposes," *European Accounting Review* 15(3):367.
- Austin, R. D. 1996. *Measuring and Managing Performance in Organizations*. New York: Dorset House Publishing.
- Bhimani, A., Gosselin, M., Ncube, M., and Okano, H. 2007. "Activity-based costing: How far have we come internationally?" *Cost Management* 21(3)(May/Jun):12.
- Bourne, M., Neely, A., Platts, K., and Mills, J. 2002. "The success and failure of performance measurement initiatives: Perceptions of participating managers," *International Journal of Operations & Production Management* 22(11):1288.
- Bragg, S. M. 2007a. *Management Accounting Best Practices: A Guide For The Professional Accountant*. Hoboken, NJ: John Wiley & Sons.
- Bragg, S. M. 2007b. *Throughput Accounting: A Guide to Constraint Management*. Hoboken, NJ: John Wiley & Sons.
- Brandt, D. 2009. "Searching for lean with lean," *Industrial Engineer* 41(5)(May):50.

- Brewton, J. 2009. "The lean office: Develop lean administrative procedures," *Cost Management* 23(2) (Mar/Apr):40.
- Brown, M. G. 2007. *Beyond the Balanced Scorecard: Improving Business Intelligence with Analytics*. New York: Productivity Press.
- Buhovac, A.R. and Slapnicar, S. 2007. "The role of balanced, strategic, cascaded and aligned performance measurement in enhancing firm performance," *Economic and Business Review for Central and South-Eastern Europe* 9(1)(Feb):47.
- Cardinaels, E. and Labro, E. 2009. "Costing systems," *Financial Management* (Dec/Jan 2008):42.
- Church, A. H. 1908. *The Proper Distribution of Expense Burden*. Baltimore, MD: Waverly Press.
- Church, A. H. 1917. *Manufacturing Costs and Accounts*. New York: McGraw-Hill Book.
- Cohen, S., Venieris, G., and Kaimenaki, E. 2005. "ABC: Adopters, supporters, deniers and unawares," *Managerial Auditing Journal* 20(8/9):981.
- Cokins, G. 2001. *Activity-Based Cost Management: An Executive's Guide*. New York: John Wiley & Sons.
- Cooper, R. 2000. "Viewpoint: 21st century cost management Cost management: From Frederick Taylor to the present," *Cost Management* 14(5) (Sep/Oct):4.
- Cooper, R., Kaplan, R. S., Maisel, L. S., Morrissey, E., and Oehm, R. M. 1992. *Implementing Activity-Based Cost*

Management: Moving from Analysis to Action.
Montvale, NJ: Institute of Management Accountants.

Cruikshank, J. L. 1987. *A Delicate Experiment: The Harvard Business School 1909-1945.* Boston, MA: Harvard Business School Press.

Cunningham, J. E. and Fiume, O. J. 2003. *Real Numbers: Management Accounting in a Lean Organization.* Durham, NC: Managing Times Press.

Dearlove, D. and Crainer, S. "Whatever happened to yesterday's bright ideas?" In The Conference Board Review [database online]. New York. Available from <http://www.conference-board.org/articles/articlehtml.cfm?ID=346>. [Accessed July 25, 2009]

Deem, J. 2009. The relationship of organizational culture to balanced scorecard effectiveness. D.B.A., Davie, FL: Nova Southeastern University.

Does, R., Vermaat, T., Verver, J., Bisgaard, S., and Van Den Heuvel, J. 2009. "Reducing start time delays in operating rooms," *Journal of Quality Technology* 41(1) (Jan):95.

Dopuch, N. and Birnberg, J. G. 1969. *Cost Accounting: Accounting Data for Management Decisions.* New York: Harcourt, Brace & World.

Dvorak, P. 2009. "Clarity is missing link in supply chain," *The Wall Street Journal*, May 18, sec A.

Eden, Y. and Ronen, B. 2007. *Approximately Right, Not Precisely Wrong: Cost Accounting, Pricing and*

Decision Making. Great Barrington, MA: North River Press.

Edwards, R. and Boyns, T. 2009. *The History of Cost and Management Accounting: The Experience of the United Kingdom*. Oxford, UK: Routledge.

Everaert, P. and Bruggeman, W. 2007. "Time-driven activity-based costing: Exploring the underlying model," *Cost Management* 21(2)(Mar/Apr):16.

Ferrara, W. L. 2007. "Topics worthy of continued discussion and effort—Even after forty years of trying," *Journal of Management Accounting Research* 19:171-179.

Fleischman, R. K. and Parker, L. D. 1990. "Managerial accounting early in the British industrial revolution: The Carron Company, a case study," *Accounting and Business Research* 20(79)(Summer):211.

Fleischman, R. K. and Parker, L. D. 1997. *What Is Past Is Prologue*. London: Garland Publishing, Inc.

Follett, M. P. and Sheldon, O. 2003. *Dynamic Administration, the Collected Papers of Mary Parker Follett: Early Sociology of Management and Organizations*. London: Routledge.

Garrison, R. H., Noreen, E. W., and Brewer, P. C. 2010. *Managerial Accounting*. 13th ed. New York: McGraw-Hill.

Geri, N. and Ronen, B. 2005. "Relevance lost: The rise and fall of activity-based costing," *Human Systems Management* 24(2):133.

- Goldratt, E. M. 1990. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Crotonon-Hudson, NY: North River Press.
- Goldratt, E. M. 1997. *Critical Chain*. Great Barrington, MA: North River Press.
- Goldratt, E. M. and Cox, J. 1984. *The Goal: A Process of Ongoing Improvement*. Croton-on-Hudson, NY: North River Press.
- Grudens, R. 1997. Henry Ford: Helped Lead American World War II Production Efforts. In HistoryNet.com [online database]. Available from <http://www.historynet.com/henry-ford-helped-lead-american-world-war-ii-production-efforts.htm>. [Accessed July 20, 2009].
- Harris, N. J. 1936. "What did we earn last month?" *National Association of Comptrollers and Accounting Officers Bulletin* XVII(5)(Jan 15):501-502.
- Hayes, R. H. and Abernathy, W. J. 1980. "Managing our way to economic decline," *Harvard Business Review* 58 (4)(July-August 1980):67.
- Henkoff, R. 1998. "Boeing's big problem," *Fortune* 137 (1)(Jan 12):96.
- Hilton, R. W. 2009. *Managerial Accounting: Creating Value in a Dynamic Business Environment*. 8th ed. New York: McGraw-Hill.
- Hope, J. and Fraser, R. 2003. *Beyond Budgeting: How Managers Can Break Free From the Annual Performance Trap*. Boston, MA: Harvard Business School Press.

- Huntzinger, J. R. 2007. *Lean Cost Management: Accounting for Lean by Establishing Flow*. Ft. Lauderdale, FL: J. Ross Publishing.
- Jargon, J. 2009. "Latest Starbucks buzzword: 'lean' Japanese techniques," *The Wall Street Journal*, August 4, sec A.
- Jelinek, M. 1980. "Toward systematic management: Alexander Hamilton Church," *Business History Review* 54(1)(Spring):63-79.
- Jenson, B., Albrecht, D., Walters, P. "Accounting education using computers and multimedia." Available from <http://pacioli.loyola.edu/aecm/>(accessed August 29, 2009).
- Johnson, H. T. and Broms, A. 2000. *Profit Beyond Measure*. New York: The Free Press.
- Johnson, H. T. and Kaplan, R. S. 1987. *Relevance Lost: The Rise and Fall of Management Accounting*. Boston, MA: Harvard Business School Press.
- Kanigel, R. 1997. *The One Best Way: Frederick Winslow Taylor and the Enigma of Efficiency*. New York: Viking Penguin.
- Kaplan, R. S. 1984. "The evolution of management accounting," *The Accounting Review* 59(3)(Jul):390.
- Kaplan, R. S. and Anderson, S. R. 2003. Time-Driven Activity-Based Costing, Working Paper # 04-045. Boston, MA: Harvard Business School. <http://www.hbs.edu/research/facpubs/workingpapers/papers2/0304/04-045.pdf>. [Accessed August 1, 2009].

- Kaplan, R. S. and Norton, D. P. 1992. "The balanced scorecard—Measures that drive performance," *Harvard Business Review* 70(1)(Jan/Feb 1992):71-79.
- Kaplan, R. S. and Norton, D. P. 1996. *The Balanced Scorecard*. Boston, MA: Harvard Business School Press.
- Kelly, J. and Rivenbark, W. 2008. "Budget theory in local government: The process-outcome conundrum," *Journal of Public Budgeting, Accounting & Financial Management* 20 (4)(Winter):457.
- Kiani, R. and Sangeladji, M. 2003. "An empirical study about the use of the ABC/ABM models by some of the fortune 500 largest industrial corporations in the USA," *Journal of American Academy of Business, Cambridge* 3(1/2)(Sep):174.
- Lawson, R., Stratton, W., and Hatch, T. 2003. "The benefits of a scorecard system: A new North American study explains how balanced scorecard users get their money's worth," *CMA Management* 77(4)(Jun/Jul):24.
- Liker, J. K. 2004. *The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer*. New York: McGraw-Hill.
- Liker, J. K. and Meier, D. 2006. *The Toyota Way Field Book: A Practical Guide for Implementing Toyota's 4Ps*. New York: McGraw-Hill.
- Litterer, J. A. 1961. "Alexander Hamilton Church and the development of modern management," *Business History Review* 35(Summer 1961):211-225.

- Maskell, B. and Baggaley, B. 2004. *Practical Lean Accounting: A Proven System for Measuring and Managing the Lean Enterprise*. New York: Productivity Press.
- McFarland, W. B. 1950. "How standard costs are being used today for control, budgeting, pricing: A survey," *Journal of Accountancy (Pre-1986)* 89(2) (Feb):125.
- McLean, T. 2006. "Continuity and change in British cost accounting development: The case of Hawthorn Leslie, shipbuilders and engineers, 1886-1914. *The British Accounting Review* 38(1)(Mar):95.
- Nolan, G. J. 2005. "The end of traditional budgeting," *Journal of Performance Management* 18 (1):27-39.
- Oliver, L. 2004. *Designing Strategic Cost Systems*. Hoboken, NJ: John Wiley & Sons.
- Olson, R., Verley, J., Santos, L., and Salas, C. 2004. "What we teach students about the Hawthorne studies: A review of content within a sample of introductory I-O and OB textbooks," *The Industrial-Organizational Psychologist* 41(3)(January):23-39.
- Palmer, R. J. and Vied, M. 1998. "ABC: Could ABC threaten the survival of your company?" *Management Accounting* 80(5)(Nov):33.
- Peters, T. J. and Waterman Jr., R. H. 1982. *In Search of Excellence: Lessons From America's Best-Run Companies*. New York: Harper & Row.

- Polischuk, T. 2009. "What's lean mean? . . ."
PackagePrinting 56(2)(Feb):30.
- Pullin, J. 2009. "The learning factory,"*Professional Engineering* 22(11)(Jun 24):31.
- Ricketts, J. A. 2008. *Reaching the Goal: How Managers Improve a Services Business Using Goldratt's Theory of Constraints*. Boston MA: IBM Press and Pearson Education, Inc.
- Schragenheim, E., Dettmer, H. W., and Patterson, J. W. 2009. *Supply Chain Management at Warp Speed*. Boca Raton, FL: Auerbach Publications, Taylor & Francis Group.
- Shewhart, W. E. 1931, 1980. *Economic Control of Quality of Manufactured Product*. New York: D. Van Nostrand Company, Inc.; Milwaukee, WI: American Society for Quality Control.
- Shipulski, M., Hockley, R., and Beck, R. 2009. "Resurrecting manufacturing,"*Industrial Engineer* 41(7)(Jul):24.
- Shook, J. 2009. "Toyota's secret,"*MIT Sloan Management Review* 50(4)(Summer):30.
- Skapinker, M. 1998. "Boeing, Boeing, bong: Michael Skapinker on the production woes that have spoiled what should have been a bumper period for the U.S. aircraft maker,"*Financial Times*, February 6.
- Smith, D. 2000. *The Measurement Nightmare*. Boca Raton, FL: St. Lucie Press.

- Speckbacher, G., Bischof, J., and Pfeiffer, T. 2003. "A descriptive analysis on the implementation of balanced scorecards in German-speaking countries," *Management Accounting Research* 14(4) (Dec):361.
- Stewart, M. 2009. *The Management Myth: Why the Experts Keep Getting It Wrong*. New York: W. W. Norton & Company, Inc.
- Stratton, W., Desroches, D., Lawson, R., and Hatch, T. 2009. "Activity-based costing: Is it still relevant?" *Management Accounting Quarterly* 10(3) (Spring):31.
- Stuart, I. and Boyle, T. 2007. "Advancing the adoption of 'lean' in Canadian SMES," *Ivey Business Journal Online* (Jan/Feb), http://www.iveybusinessjournal.com/article.asp?intArticle_ID=650 [Accessed April 8, 2010].
- Taylor, F. W. 1911, 1967. *The Principles of Scientific Management*. New York: Harper & Row; W.W. Norton & Company, Inc.
- Tyson, T. 1993. "Keeping the record straight: Foucauldian revisionism and nineteenth century U.S. cost accounting history." *Accounting, Auditing & Accountability Journal* 6(2):4.
- Van Veen-Dirks, P. and Molenaar, R. 2009. "Customer profitability pricing," *Cost Management* 23(3) (May/Jun):32.
- Vangermeersch, R. and Schwarzback, H. R. 2005. "The historical development of management accounting."

In Weil, R. L. and Maher, M. W., eds., *Handbook of Cost Management*. 2nd ed. Hoboken, NJ: John Wiley & Sons.

Weber, J. and Linder, S. 2005. "Budgeting, better budgeting, or beyond budgeting," *Cost Management* 19(2)(Mar/Apr):20.

Weil, N. 2007. "A legacy of failure; researchers cite a 90 percent failure rate among companies trying to execute their strategies. What's up with that?" *CIO* 20(19)(Jul 15):1.

Whitehead, T. N. 1938. *The Industrial Worker: A Statistical Study of Human Relations in a Group of Manual Workers*, Volumes I & II. Cambridge, MA: Harvard University Press.

Womack, J. P., Jones, D.T., and Roos, D. 1990. *The Machine That Changed the World*. New York: Rawson Associates.

About the Author

Charlene Spoede Budd is a Professor Emeritus from Baylor University, where she taught management accounting and project management classes for a number of years. She received her undergraduate degree (accounting major, *Summa Cum Laude*), and MBA degree from Baylor University (1972 and 1973, respectively) and her PhD from The University of Texas at Austin (1982), where she specialized in the fields of accounting, economics, and finance. She holds the following active professional designations: CPA; CMA, CFM, PMP. In addition, she is certified in all areas of the Theory of Constraints by the Theory of Constraints International Certification

Organization (TOCICO). Her research has been published primarily in practitioner journals and she has been awarded three Certificates of Merit for articles published in *Strategic Finance*. She also has singly or coauthored publications in *Industrial Marketing Management* (special issue on projects), *Human Systems Management Journal*, *Today's CPA*, *The Counselor*, and other journals and many conference proceedings. Dr. Budd has coauthored two accounting textbooks and she and current coauthor, Charles Budd, have published *A Practical Guide to Earned Value Project Management* (Management Concepts, 2005 and 2010) and *Internal Control and Improvement Initiatives* (BNA, 2007). She is active in several professional organizations, including the American Accounting Association, Financial Executives Institute, and Project Management Institute. In addition, she has been a member of the AICPA's Content Committee and was Chair of the Business Environment and Content Subcommittee of the AICPA from 2004 until 2008. Currently, she is Chair of the Finance and Metrics Committee of the TOCICO. Most of her time now is devoted to research, but she also is a member of the Board of Directors of a public company.

CHAPTER 14

Resolving Measurement/Performance Dilemmas

Debra Smith and Jeff Herman

Introduction

What are *measurement/performance dilemmas*? For the purposes of this chapter, let's say that they are *situations that pull people, departments, divisions, and companies in opposite or competing directions*. For example, it is the purchasing agent who is torn between selecting the lowest cost supplier versus selecting the most reliable supplier; the shift supervisor who waffles on whether to authorize overtime; the salesperson who pleads for an earlier commitment date versus the scheduler who doesn't want to disrupt the schedule; the controller who wants to outsource versus the plant manager who wants to keep the business in-house; the CFO who wants to slash inventories versus the vice president of sales who wants to maintain or even increase inventory; and the engineering manager who wants to standardize products versus the sales manager who wants to sell customized solutions. These dilemmas often represent a constantly changing and frustrating series of daily, weekly, quarterly, and annual set of unsatisfactory compromises. These compromises can cost the organizations tremendous amounts of money as people and resources are whipsawed between two often extreme positions.

What is behind these dilemmas? Commonly these “extreme positions” represent the most apparent or obvious way to meet a particular metric. In our examples, it is purchase price variance versus material availability; overtime budget versus on-time performance; booked business versus schedule stability; product cost versus volume (which can connect back to product cost); cash versus availability/fill rates; utilization versus new business development.

Are the metrics always in conflict? Of course not but often they are. In most for-profit companies, the goal takes some form of *return on investment* (ROI) or *return on average capital employed* (RACE). The strategy to accomplish this goal almost always includes tactical objectives ([Fig. 14-1](#)) to:

1. Decrease inventory
2. Improve quality
3. Increase sales

Copyright © 2010 by Debra Smith and Jeff Herman.



FIGURE 14-1 Tactical objectives to increase ROI.

4. Decrease cost
5. Improve due date performance (DDP)

Management assumes that improving these five tactical objectives will drive ROI in the right direction. Their assumption is absolutely valid. The problem is that often the metrics and corresponding actions to achieve these seemingly straightforward tactical objectives will, and do, constantly come into conflict with each other.

Often when a company grows to a relatively modest size, it becomes necessary to segment the organization into areas of functional responsibility (i.e., Sales, Manufacturing, Finance, etc). Therefore it is logical that the tactical objectives are assigned to the functional managers to focus on and improve. However, can a drive to increase quality drive costs up and increase cycle time? Can a drive to decrease costs negatively impact quality and our marketplace? Can a drive to increase sales erode margins? Can a drive to increase on-time delivery or shorten our lead time, increase costs and inventory and erode quality? Can programs to decrease inventory starve the plant and result in decreased on-time delivery and increased overtime costs as well as increased cycle time and *work-in-progress* (WIP)? In reality the answer to these questions is, "YES!" Each local manager, measured on improving his or her functional responsibility, will drive the organization directly into conflict with itself. This, by definition, is extremely wasteful and limits the organization from having any type of dramatic improvement.

Is there no solution? For years, companies that have embraced the *Theory of Constraints* (TOC) have proven that there is indeed a solution. When properly aligned in a TOC system, moving all of the objectives in the right direction simultaneously and without conflict is achievable. [Figure 14-2](#) shows the results of a review of the literature by Mabin and Balderstone (2000) of 82 TOC case studies from around the globe. This review showed that companies that implemented TOC were able to move these tactical objectives simultaneously in the right direction.

Do We Measure Too Much?

Once again, are metrics always in conflict? No, but often they are and when they are not in conflict the assumptions around how to achieve certain metrics can put people in conflict. One conclusion we can draw from this is that the more metrics an organization has, the more potential there is for those metrics or the assumptions of how to achieve those metrics to be in conflict.

Modern corporations have metrics everywhere and they devote tremendous amounts of resources and energy to maintaining them. What is interesting is that the number of measures, like the universe, always seems to be expanding (even accelerating). An analogy (with a direct connection to this topic, by the way) is with modern *Enterprise Resources Planning* (ERP) systems. Ask any ERP provider how many lines of code they had 10 years ago versus how many they have today (they may not even be able to give you a number). The irony or perhaps the lesson is that most of their customers will candidly admit (behind closed doors) that those systems have not really produced any better business results over that 10-year period; they are just more costly to operate. Are we making it harder than it needs to be?

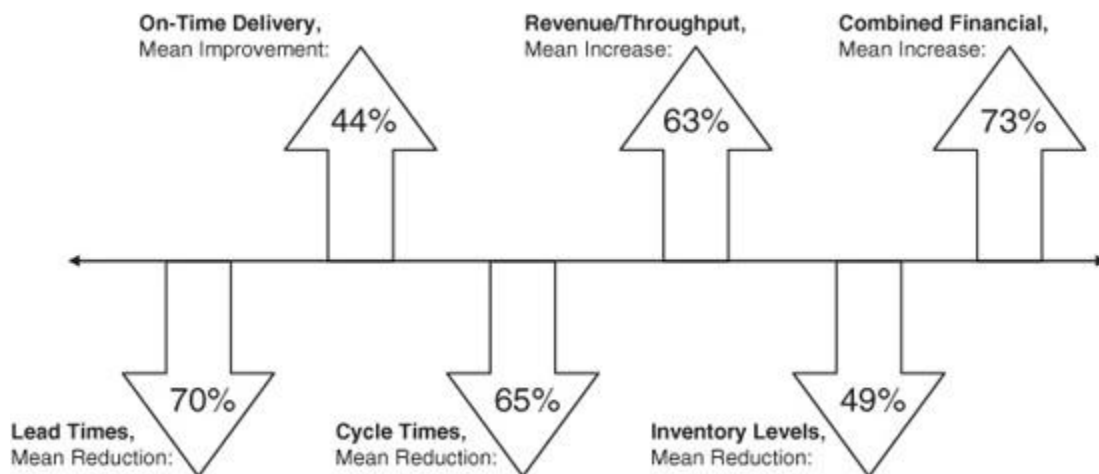


FIGURE 14-2 Average improvement in measures of companies after implementation of TOC. (*From Mabin, V. and Balderstone, S. 2000. The World of Theory of Constraints. Boca Raton, FL: St. Lucie Press.*)

Maybe in trying to control everywhere we end up controlling nowhere. Albert Einstein once said, “Any intelligent fool can make things bigger and more complex, it takes a touch of genius and a lot of courage to move in the opposite direction.” He also said, “Everything should be made as simple as possible and not simpler.” In these two statements, he cleverly lays out the criteria for effective problem solving and control. Solutions should be elegant, meaning concise and simple, but at the same time, all truly relevant factors must be considered. This is the direction of the solution for resolving measurement/performance dilemmas.

Ultimately, what is needed are measurements that contain a set of relatively simple, highly visible execution priorities to focus and align the entire team of functional managers around actions that have the greatest organizational ROI regardless of the impact on the tactical objectives. In other words, all of the objectives should be understood relative to their current impact on ROI.

The achievement of this solution will reduce the number of primary metrics and the corresponding potential for conflicts between metrics as well as better clarify the actions needed to meet those metrics.

Why Do We Have Measurements?

The point of any system of measurement should be to:

- Judge progression toward a specific goal or objective.
- Drive behavior toward a specific goal or objective.

- Highlight relevant factors in relation to achieving the goal or objective.

In a recent *USA Today* piece by Bruce Horowitz (2009), Douglas Conant, President and CEO of Campbell's Soup said, "You can't talk your way out of something you behaved your way into." Assuming an organization has a defined objective, one of the keys to reaching that objective (in any time frame) is to get all of the components of the system to behave in a manner that moves the system toward the objective. Why is there a link between measurements and behavior? The saying, "tell me how you measure me and I will tell you how I behave," has always been a cliché linking behavior to measurements. It is a cliché because, while true, it is an oversimplification of the relationship between metrics and behavior. Not only is it simply the existence of a metric that drives a specific behavior, it is also the lack of another directly conflicting metric that drives directly conflicting behavior under the same circumstances and a system of feedback and accountability that removes the cliché label from the phrase.

This means that metrics must be coordinated and constructed in order to induce local areas to work together to do what is in the interest of the whole and those metrics must be backed up by a robust system of accountability and visibility (that in itself needs to be measured). This is a basic building block to organizational synchronization and efficiency. If resources are not behaving in a synchronized manner, then some conclusions might be drawn:

1. Formal metrics are not in synchronization. There are potential conflicts or disconnects in the formal metric system. As mentioned previously, the more metrics a

system has in place, the greater the chance of conflict.

2. There are no formal metrics or there are significant gaps in the formal metric system. This means resources tend to drive behavior around how they perceive they are measured or believe they should be measured. In the absence of a formal metric, this perception is often driven by a resource's own view of what the right thing is. This creates the opportunity for conflicts driven by interpretation or assumptions.
3. There are formal metrics, there are no conflicts in those metrics, and there are no significant gaps between those metrics but there is no effective feedback and accountability system. Many of us can probably remember coworkers who did whatever they wanted to do regardless of what they probably should do with little or no individual consequences. Additionally, many of us can remember a situation in which behaviors persisted according to a metric that was obsolete. Why was the metric still in place? There was no effective feedback system to point out that it needed to be changed or eliminated.

The question becomes, how do we set up a formal and coordinated system of metrics without significant gaps and conflicts with clear feedback and accountability?

This chapter is organized into three sections. The first section explores the basic global metric dashboard that companies should use to facilitate and judge progression in relation to the goal. The second section outlines the basic coordinated local measure dashboard that supports the global measures. The final section lays out how to build an effective feedback system to drive visibility and accountability in order to better resolve any remaining dilemmas and drive continuous improvement.

Global Metrics

This chapter assumes that a company has a defined goal and strategy. If there is no defined goal and strategy, then why measure anything? See Chapters 17, 18, and 19 in this Handbook for company goals and strategy.

Critical performance measures at the global level ultimately boil down to one basic measure of performance; some form of measure for return on equity. The specific return equation that a company uses is essentially irrelevant. Common measures are ROI, RACE, *return on capital employed* (ROCE), and *return on assets* (ROA). Essentially, two components come together to create that equation in whatever forms the company has chosen to use:

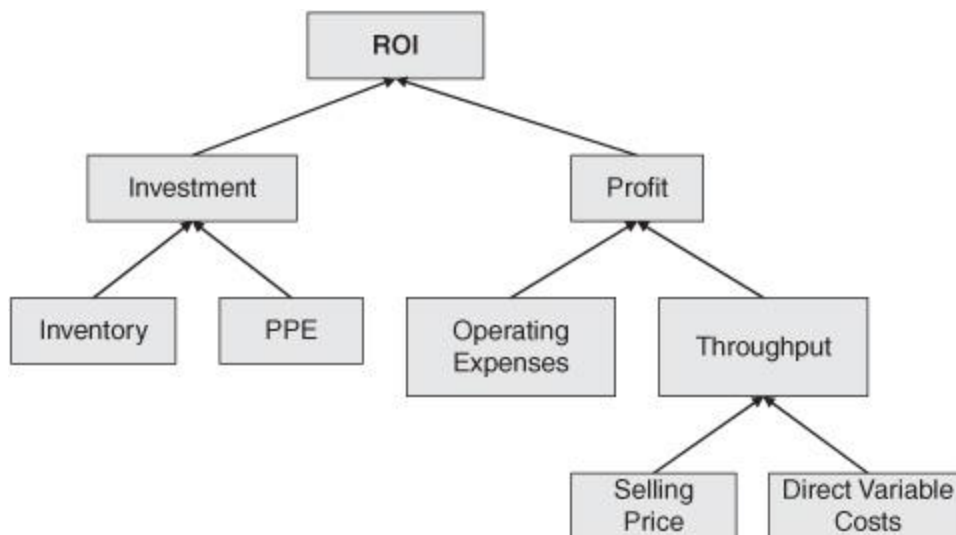


FIGURE 14-3 The components of ROI.

- A measure of profit. *Profit can be derived simply by the equation of Throughput (T) minus Operating Expense (OE). Throughput is calculated both at the aggregate and product levels by taking sales dollars minus all*

direct variable costs. Direct variable cost (also called totally variable cost) is any expense that has a one-for-one direct relationship to the product or service, raw materials, freight, sales commissions, etc. OEs are all of the expenses of a business other than the directly variable cost of the product. This accounting approach eliminates the distortion in earnings between periods when the product produced is greater or less than the product sold by eliminating the allocation of fixed costs to inventory. This approach does not reward the building of inventory that does not protect either current or seasonal future Throughput. This approach better aligns cash flow of the period with the income statement of the period.¹

- A measure of investment or capital employed. *Capital employed* has many definitions. *In general, it is the capital investment necessary for a business to function. It is commonly represented as total assets less current liabilities or fixed assets plus working capital.* In most companies, there are two predominant factors in the *investment* equation. The first is the total amount of *inventory*. The second is commonly referred to as *Property, Plant, and Equipment (PP&E)*.

[Figure 14-3](#) shows the global metric hierarchy.

Effectively deploying these metrics, however, requires certain relevant factors to be understood and defined before any company's system of metrics can produce the meaningful and relevant information for good decision-making and a reduction in the number of measurement-related dilemmas.

Measurement is largely executed to judge performance. Unfortunately, even a judgment as seemingly straightforward as profitability could be disastrously flawed.

One example is the common valuation of inventory as it relates to profitability within measurement periods. For decades now, we have known about the negative consequences associated with too much inventory—increased costs, decreased flow, damaged goods, etc. However, how do we judge it on the balance sheet? Not only do we value it as an asset, but we “add value” to it as we absorb labor and overhead into the inventory. This means that companies can build inventories and declare a profit without the sales to support it.² This means that a dramatic attempt to decrease inventory for improvement in flow and Throughput can result in the punishment of these “go getters” for poor short-term profitability performance or low absorption rates. Yet, at the same time, inventory tends to be a critical measurement within measurement periods as well. Trying to balance these competing factors can lead to absurd behavior. We worked with a company a few years back that would refuse the receipt of incoming inventory at the end of every month (a measurement period) only to expedite it a week later and work overtime to attempt to meet on-time shipments. What impact does this have on cost? What is the impact on on-time delivery performance? They knew this behavior was painful to the organization and yet felt their hands were tied by the corporate measurement. Ironically, when you sit down with a CFO or Controller and explain the cause-and-effect logic of what is happening they often scream, “What!? That’s not what we want them to do!” Is this a situation of conflicting measures? Yes. Is it a poor interpretation? Probably. Does it demonstrate a lack of an effective feedback and accountability system? Unquestionably yes.

The Constraint Is the Primary Relevant Factor

There is a single factor determining the flow through production and to the market. This same factor gives us the assumptions that underlie the cost and revenue opportunity of any potential action or investment. That factor is the constraint. The constraint could be a resource, raw material or purchased part, a skill set, a policy or procedure, a measure, etc. Information about the constraint is what is relevant. Information on the impact that any option has on the constraint is critical to a measurement system and it should point to actions that, when taken, will provide a bottom-line return.

Constraints change how we judge product profitability, short-term profit maximization, ROI, capital, inventory, and manpower. Constraints impact the rate at which an organization can make money—they are a system's leverage point. This is why having a TOC logistical system and its associated measurements is so vital for operating decisions and improvement in a logistics environment.

TOC is a methodology and a set of processes to maximize a system's ROI/RACE equation by employing solutions that identify, exploit, and manage a system through its leverage points and their interactions with each other.

When considering any alternative actions, plans, or improvement projects:

1. Consider the impact on the constraint's performance. This includes knowing if the constraint will move because of your decision and if so, where and what are the implications?
2. For every investment we must know how the economic return will be generated. Will the market buy more products or will there be any reduction in investment (e.g., strategic inventory buffers) or OE?

Cash inflow versus cash outflow should be one of the primary parameters of decision making. The true implications of any assumptions of the system should be judged against “cash in” and “cash out.”

Using the simple performance measures defined previously and removing all other performance metrics prevents companies from making poor decisions³ driven by performance metrics heavily skewed by rolling up labor and overhead into the “cost” of a product. For example, most measures of unit cost would suggest that any reduction in run time or setup time at any resource would result in a lower cost product, or an improvement. TOC dictates that this logic is categorically untrue. Given that this type of localized “cost” thinking is embedded in the vast majority of decision-making tools, the savings or profits generated for most improvement projects are a mirage and never materialize.

Often appropriation requests are constructed to justify investment heavily weighted on reducing cost of the product. For example, if we can reduce the time it takes in one step of the process by 25 percent, then the product cost is reduced (due to less total labor content) and this will translate to an assumed bottom-line improvement. In reality, since these investments and process improvement initiatives (driven by the product cost equation) do not consider constraints or buffers, there is no effective way to judge whether any time reductions at any resource will result in bottom-line improvement. It is quite likely that without the consideration of constraints and buffers, most changes would not result in a net positive ROI position—and often negatively impact performance of the constraint and thus of the whole system as many of these changes require an investment or spending of some sort.

If we push for cycle time improvement at a non-constraint, generally what is the predictable outcome? Remember, a

non-constraint means that this part of the system is not currently dictating the pace at which the company is making money. This means that if a non-constraint local resource were enabled to produce faster, the system would experience:

1. No increase in sales or shipments to the customer—no increase in Throughput.
2. A likely increase in parts produced that are not able to be immediately consumed—increased inventory. Increased WIP could also result in longer lead times, decreasing DDP and ultimately reducing Throughput.
3. Some investment was probably made to make the improvement. Potentially, there are additional space requirements or costs of borrowing associated with increased Inventory. In addition, most often the improved rate of production did not allow for any reductions in labor—no decrease in OE.

In other words, locally judging this “improvement” results in Throughput, Inventory, and OE moving in the wrong direction. Judging the potential action on its impact on the constraint would have resulted in saving this money in investment for something that would provide the opportunity for a real bottom-line return.

If one thing can be learned from this book it is the understanding of the impact the constraint has on the system. If the 25 percent improvement in velocity happens to be on a bottlenecked resource, the impact to the bottom line would likely be *much* greater than the small cost savings that the product cost measurement suggests. In addition to the savings, increased Throughput would result thus improving the organization’s bottom-line. Thus, the product cost argument would dramatically understate the need for this improvement.

The discussion above helps explain why so often cost reduction projects approved by senior managers end up having no impact on the bottom line at all, and then why management have such difficulty understanding what went wrong.

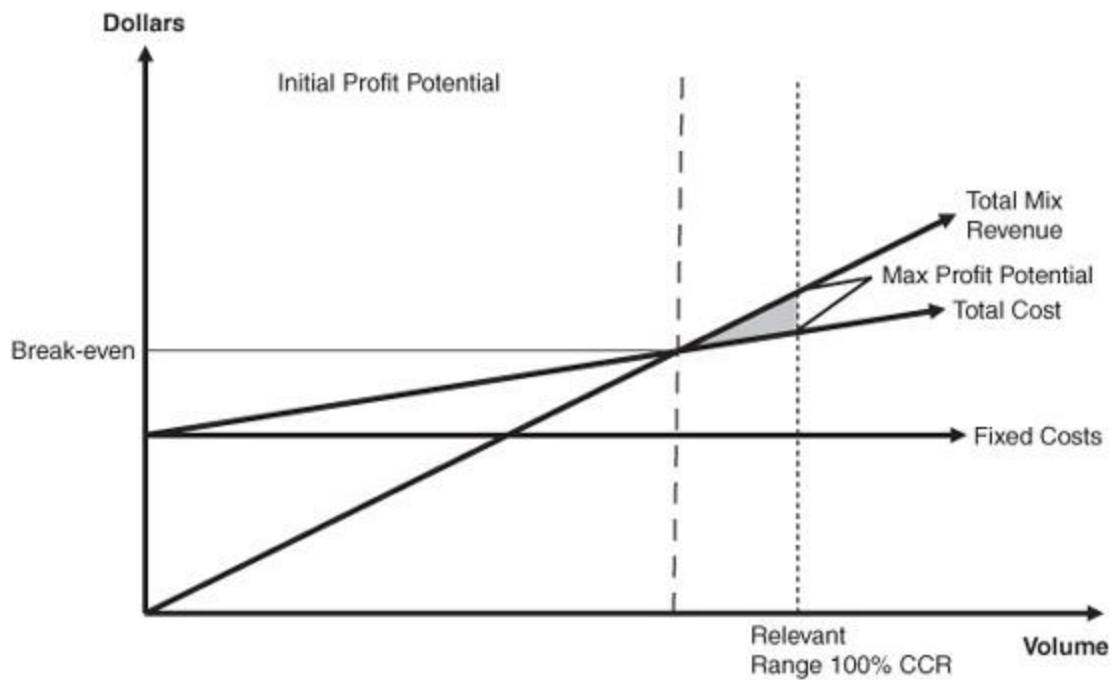


FIGURE 14-4a TOC break-even chart of initial profit potential.

Profit Maximizing in TOC

Since most for-profit companies have a goal related to ROI or RACE and profit is a major component of those measures, we need to understand the basic strategy for profit maximization for each company. Remember one of the most relevant factors is the location of the constraint, the company's major leverage point. TOC goes back to fundamental economics as the basis for management accounting information (Horngren et al., 1993, 156) to maximize profit as demonstrated by a quote from a popular management accounting text:

The criterion for maximizing profits when one factor limits sales is to obtain the greatest possible contribution to profit for each unit of the limiting or scarce factor. The product that is most profitable when one particular factor limits sales may be the least profitable if a different factor restricts sales. When there are limitations, the conventional contribution or gross margin per sales dollar ratios provide an insufficient clue to profitability.

[Figure 14-4a](#) demonstrates the cost and revenue potential of a system through a simple break-even chart.⁴ Note that the “fixed costs” in the diagram include all OE as defined previously. The “total cost” line is all of the direct variable costs that are added on top of the fixed cost baseline that is associated to the sale of product. This company’s revenue potential is determined by the intersection of the relevant range and the total mix revenue. The company profit potential is the revenue potential minus the total cost at any given point above the break-even point.

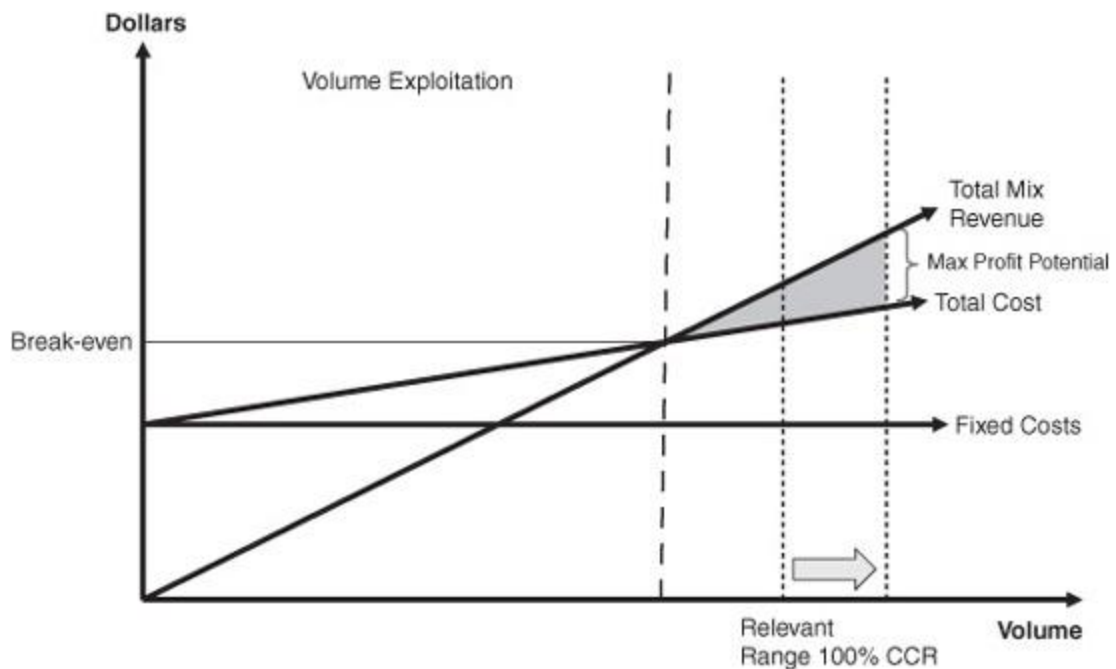


FIGURE 14-4b TOC break-even chart of volume exploitation.

If top management's attention is focused on investments directly impacting the bottom line as opposed to being distracted by endless requests in the name of localized improvement, then relatively rapid and significant organization improvement is close at hand.

The limitation of the constraint is *the* key factor that determines overall capacity and therefore the initial relevant range potential in [Fig. 14-4a](#). By identifying and exploiting the constraint—the organization's leverage point—and by applying all brain power and focus to squeezing more out provides a great opportunity to have an immediate and long-term impact to the bottom line for minimal or no investment.

Exploiting the constraint has two levels. *The first level is first and foremost about increasing volume.* This increase in volume can happen from two primary avenues, both of which *require* knowledge of the position and status of the constraint or drum. The first avenue is through volume that can be increased by squeezing more out of the constraint itself. This can be accomplished through a number of methods including improvements in its run rate, eliminating its starvation, or minimizing or reducing its setup time. The second avenue that volume can increase is through driving the sale of *free products*—a product that is free from passing through the constraint. Free product volume must be carefully managed so that it does not create an additional constraint. Careful management often means some sort of governing mechanism to adjust volume in relation to the total system's effectiveness to support the constraint. If there is too much free product volume, it can often mean that the resources involved in making it will have less overall *sprint or protective capacity*⁵. This means that they are less responsive to the constraint or (once downstream from the constraint) the customer. The danger of this is obvious. It can cause disruptions to the constraint,

late shipments, expedites or bigger buffers (time or stock) impacting lead times, DDP, or cash in inventory. [Figure 14-4b](#) shows the impact of volume-based exploitation techniques. Increasing the volume has expanded the relevant range of the system, which translates to higher total potential revenue and profit.

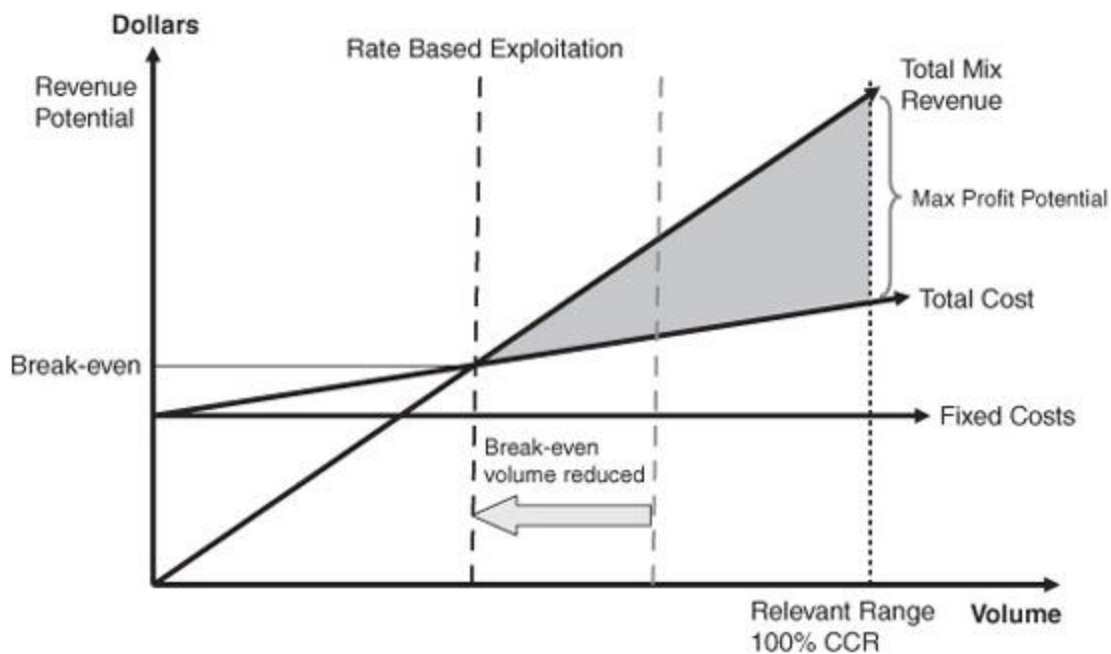


FIGURE 14-5 TOC break-even chart of rate-based exploitation of the resource constraint.

The second level of exploitation is about rate. Now that capacity/volume has been maximized, decisions must be made about which products create more profit relative to that available capacity/volume. The primary metric here is the rate at which products generate Throughput across the constraint. Using our clients as a benchmark over the past 15 years, we have seen the rates of Throughput generation per unit of constraint time by product differ by as little as \$3 to \$1 and as great as \$20 to \$1. Of course, free products do not cut across the defined constraint and thus their relative rate in comparison to each other is simply the calculated

Throughput margin (Selling Price – Direct Variable Costs) per unit of product. Notice this “new” company in [Fig. 14-5](#) has a much lower break-even point and is less at risk in a downturn as well as in the best position to capitalize on a market upturn. This happens because constraint productivity is improved by product selection of the highest Throughput per constraint unit in defining product mix, resulting in an upward thrust in the total mix revenue as seen in the steeper slope of the total mix revenue line in [Fig. 14-5](#).

Of course, these rates can be a moving target as prevailing prices and the costs of material and components can change, which in turn can change the Throughput per unit of product. This requires the use of a *pricing indifference model*. *A pricing indifference model is a tool to show at what point companies become indifferent to which product the limited capacity will be dedicated to producing, for example, product A versus product B as relevant factors for each change. Relevant factors include any significant changes to Throughput rates or capacities.* Many companies use a targeted aggregate Throughput rate in the annual budgetary process and judge progress and action around maintaining or exceeding that rate.⁶

These exploitation techniques provide a simple and level playing field (in replacement of traditional product cost and margin) to assess products against each other on the rate at which they generate cash, the selling price that different products need to capture, and the proper mix to seek in the market for the newly exposed capacity derived from constraint exploitation and free product emphasis. This global performance metric, ROI, helps focus management on the fundamental factors that influence the company’s goal and strategy and significantly reduces the number of management dilemmas.

Local Metrics

Once again, metrics need to encourage the right behavior. When dealing with an organization of size and complexity, it always seems to be a challenge to construct a system of local metrics that:

- Encourages the local parts to do what is in the interest of the global objective.
- Provides relatively clear conflict resolution between and within the local parts.
- Provides clear and visible signals to management about local progress and status relative to the organizational objectives.

A relatively simple set of six general measurements for localities is given. It is important to note that these local metrics assume that a valid TOC model has been implemented.

1. Reliability
2. Stability
3. Speed/Velocity
4. Strategic Contribution
5. Local OE
6. Local Improvement/Waste

Depending on the organization and the functional responsibilities of each organization, these metrics will be translated to very specific forms. Due to space limitations,

our examples of specific metrics will be oriented toward Operations only.

Metric 1: Reliability

The objective of this metric is to measure execution compliance to a plan or schedule. When localities (resources, work centers, processes, departments, etc.) and systems are less reliable, it requires systems to hold excessive buffer positions (time, stock, or capacity). Time,⁷ stock,⁸ and capacity⁹ are all interchangeable investments in production capacity. All three are simply stored time. Conversely, when localities can reliably perform within planned time horizons, it reduces the amount of buffer required. This reliability is pivotal to moving global metrics in the right direction. In TOC, reliability metrics are easily implemented by tracking service levels. There are obvious types of service level metrics that are important. Conventional metrics like on-time delivery and fill rates are still very important and relevant. In TOC, however, other critical service level metrics must be installed and tracked. These metrics are *performance to time* and *stock buffers*. Remember that resources feed buffers. If those resources are more reliable, it often means that buffers can be reduced. The reduction of buffers is a critical improvement objective in any TOC system.

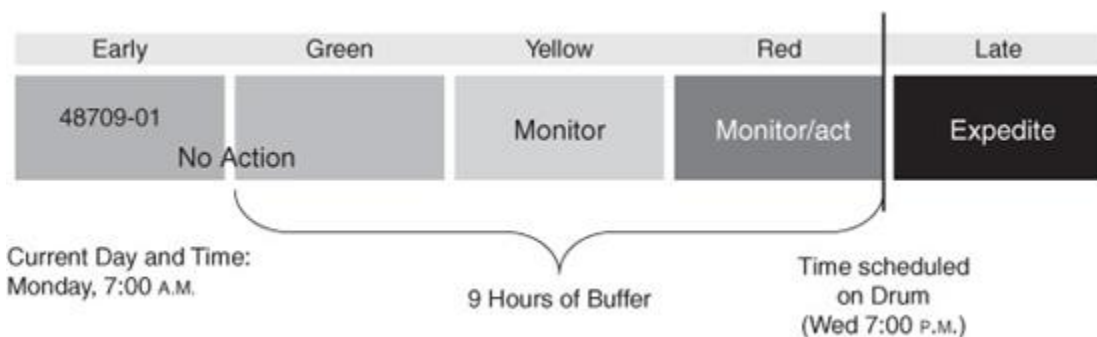


FIGURE 14-6a Stratification of a time buffer.

With regard to time buffers, this commonly means that early, expedite, and late zone penetrations into the time buffer are noted and managed at two levels—first, to direct execution actions to keep the constraint and delivery schedule stable; and second, to capture information that identifies the source of variation for future improvement activities to increase system stability. [Figure 14-6a](#) shows the stratification of a time buffer into different zones ranging from an early (far left, sometimes referred to as the light blue (LB) or white buffer zone) to expedite (red), to late delivery (far right, sometimes referred to as the dark red (DR) or black zone). The general zone color designations are provided in the figure. Notice there is a released work order (48709-01) that has not entered the time horizon that the buffer represents in front of the drum. One of these work orders will enter 9 hours out. If this facility works 24 hours a day, then that entry will occur at 10:00 A.M. on Wednesday.

We will have to reconcile the work order's actual presence in the buffer by recording when it entered the buffer and judge that against its scheduled entry to create a view about what, if any, corrective actions need to be taken. When a work order is not ready and in the buffer at the start of the green zone (scheduled buffer entry), a penetration is created in the buffer. This hole can be caused, for example, by missing materials, tooling, specs, etc. The severity of this penetration will ultimately determine when we have to act and the priority of work orders on which we have to act.

This means that we have to think about the five zones from two perspectives, "Yet to Be Received" (at the drum) and "Received" (at the drum). These are the two situations that can occur. When something is "Yet to be received" the clock is still ticking on the time it has to travel to the constraint. When something has been "Received," the hole has been filled. In [Fig. 14-6b](#), a real-time buffer board that

reconciles a released work order against their buffer status is shown.

Notice that when we account for the same time horizon from the two different perspectives, it actually creates 10 status zones. Those zones are:

1. Early—Yet to Be Received (LB). This zone actually represents all released work orders that are on the way to the buffer.
2. Green—Yet to Be Received (G). This is a hole in the buffer. Not a serious hole, but a hole nonetheless.

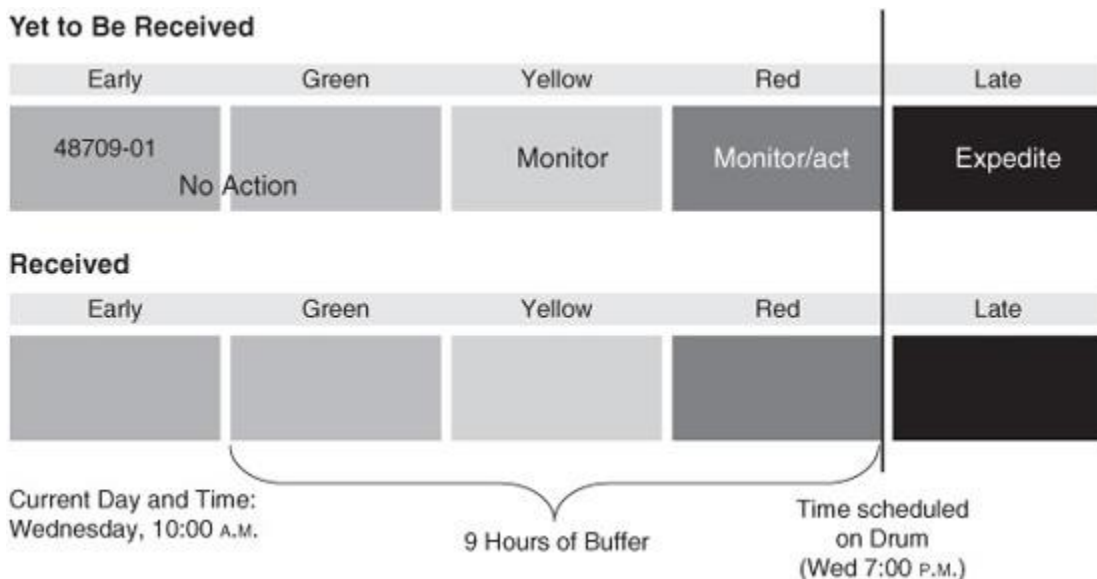


FIGURE 14-6b Reconciling released orders against buffer status.

3. Yellow—Yet to Be Received (Y). This is a deeper hole in the buffer that should now be getting the attention of the personnel responsible for managing the buffer.
4. Red—Yet to Be Received (R). This is the deepest hole that we can dig without affecting the drum schedule.

This zone alerts the appropriate personnel that if corrective actions are not taken, the drum schedule will be disrupted.

5. Late—Yet to Be Received (DR). The drum schedule has already been disrupted by this work order and it is still not present.
6. Early—Received (LB). The work order is physically present at the buffer resource and ready to be worked on by the drum ahead of the time horizon for which we are scheduled. This usually means that the standards we are using to generate the schedule may be over-estimated (very common since most companies' standards are highly inflated to try to combat Murphy and disruptions everywhere) or the work order was released ahead of schedule.
7. Green—Received (G). The work order was received within the scheduled time horizon with a relatively large amount of time to spare.
8. Yellow—Received (Y). The work order was received within the scheduled time horizon with moderate time to spare.
9. Red—Received (R). The work order was received at the constraint resource within the scheduled time horizon with little time remaining before it is scheduled on the constraint.
10. Late—Received (DR). The work order was received after the time it was scheduled on the drum. By definition, it has caused a disruption to the drum schedule.

In *Drum-Buffer-Rope* (DBR), “Early,” “Late,” and “Red” zone arrivals should require a reason code to be attached to

the job explaining *why* this work arrived when it did. In the previous example of a buffer board, the reason code is forced in order for a work order to move from the “Yet to Be Received” status to the “Received” status. A red zone arrival is not necessarily a negative thing—in fact a good system should have approximately 20 percent of its work arriving in this zone—as it is pointing us to the work center with the greatest opportunity to apply improvement focus (i.e. Lean tools) to enable shrinking the buffer and cycle time (more on this in the following section “Metric 2: Local Improvement/Waste”).

The reason codes for early and late are essential in removing the variation from inaccurate standards and routings resulting in a more accurate model for scheduling. The beauty of TOC is that it allows any company to start on a process improvement path regardless of the state of accuracy of their routings and standards. Buffers are initially sized to absorb the system’s current variation. Entry into the received status of the buffer parts that have inaccurate routings or standards will fall outside the green and yellow zones. These part work orders will be captured in the red, late and early zones with a reason code denoting the standard is wrong or the routing is wrong. This allows a systematic method to correct those parts and remove variation. Ultimately, this allows for more accurate ropes, smaller buffers, and shorter cycle times.

Stock buffers according to the TOCICO Dictionary (Sullivan et al., 2007, 43) are defined as “(a) quantity of physical inventory held in the system to protect the system’s throughput.” (© TOCICO 2007, used by permission, all rights reserved.) In TOC, these stock buffers also have five zones for management and measurement. [Figure 14-7a](#) visibly depicts the typical stock buffer zones.

As you can see, light blue (some authors call this the white zone) depicts a position that is overstocked. Green indicates a position with ample stock, which requires no

action. Yellow indicates a stock position that is in its rebuild zone. Red typically means danger or expedite, while dark red gives a visible signal of out of stock (some authors refer to this as the black zone). The total number of parts as well as the total number of days those parts have spent in red “stocked out” and “stocked out with demand” can easily be tracked over time. An example is illustrated in Figs. 14-7b and 14-7c. In both figures, the vertical axis simply represents various part numbers. In [Fig. 14-7b](#), you can see that part 78df has been stocked out 58 days over a 180-day period. Within that 58-day period, for 34 of those days, the part has been stocked out with demand against it (represented by the dark red portion on the right side of the bar). Obviously, it is more damaging to be stocked out with demand. In [Fig. 14-7c](#), you can see that part r643 has been over the limit of the green zone for 30 days over a 180-day horizon. This kind of clear visibility dramatically increases the reliability of a materials/inventory system over conventional tools like *material requirements planning* (MRP).

The expected behaviors of reliability-based metrics are quite simple. First, localities will perform work in an accurately prioritized sequence since buffer status is a direct reflection of that priority. Second, localities are encouraged to make or buy only what is necessary relative to the buffers. Since these buffers are highly visible, there tends to be little to no conflict about what the real priority is. This is effective Buffer Management (BM). Of course, this assumes that the buffers are set up properly. For more on setting up buffers properly (including placement and sizing), see *Drum-Buffer-Rope, Buffer Management, and Distribution* of the handbook.

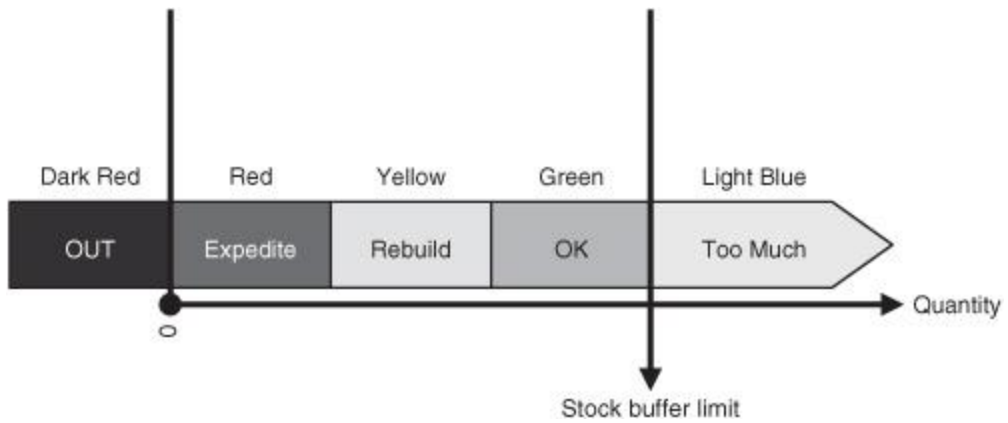


FIGURE 14-7a Stock buffer zones.

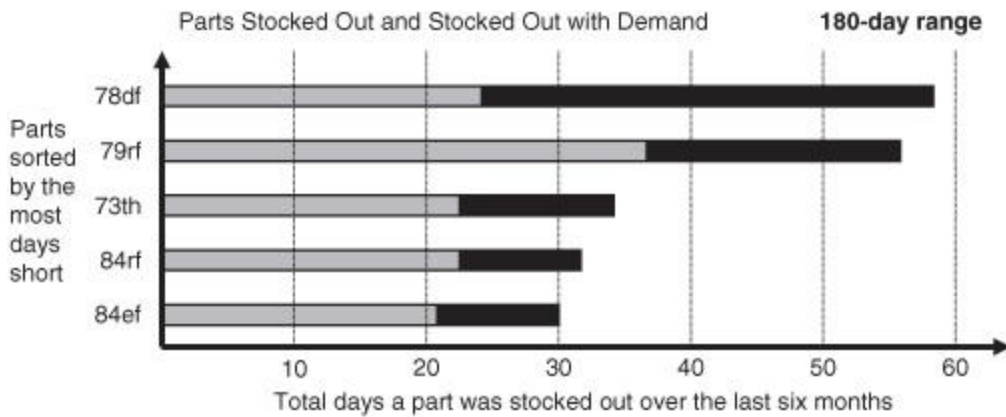


FIGURE 14-7b Number of “stockouts” and “stockouts with demand” occurrences for parts over the past 180-day period.

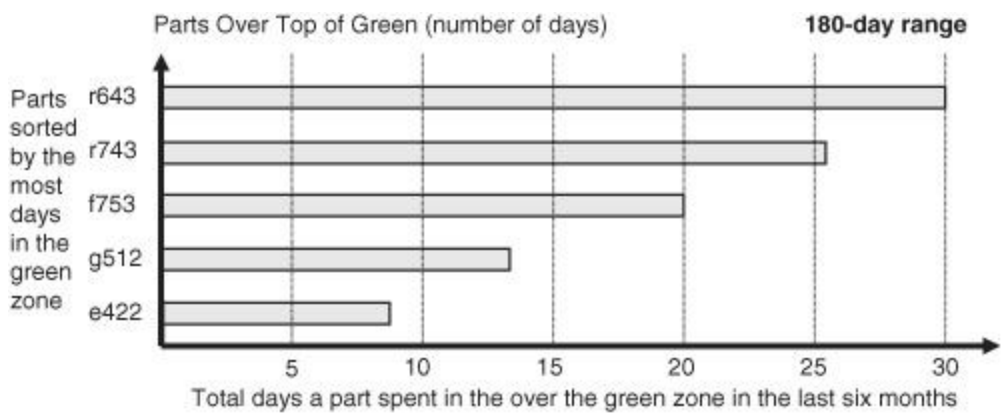


FIGURE 14-7c Number of “parts over green limit” occurrences for parts over the past 180-day period.

Metric 2: Stability

The objective of this metric is to measure the amount of variation that is passed along through the system. A key factor in overall system performance is the amount of variability and volatility that the system experiences and how well that system absorbs or deflects it away from critical areas. In particular, these critical areas are the drums in TOC systems. Encouraging stability at drums is a must. Drums are the anchor point of an overall scheduling system, meaning that all other schedules are planned from the drum schedules. If this is the case, then obviously disrupting the drum schedule creates the effect that all other schedules are out of synchronization with what is deemed to be critical. Disruptions to the drum schedules can also erode their capacity. *Drum utilization* is defined as a measure (expressed as a percentage) of how intensively the constraint resource is being used to produce Throughput. Utilization compares actual time used to produce Throughput (setup and run time) to available time of the constraint (clock time). Utilization is 100 percent minus the percent of time lost due to the constraint starvation, blockage, and breakdowns. It is critical to measure to know what the overall potential of the system is (see the “Profit Maximization in TOC” section of this chapter) and what a company is leaving on the table every measurement period. This is a dramatically different focus than traditional accounting, which has no mechanism to measure lost opportunity. In reality, there are only a few reasons that cause us to lose potential at drums:

1. Starvation. Starvation occurs when the drum runs out of material on which to work.
2. Unnecessary or over-production. This is a waste of drum capacity on things that, quite simply, are not yet required.
3. Downtime. This is downtime of the drum due to unplanned (Murphy) or planned events.
4. Blockages. Blockages occur when the drum is prevented from running because an operation that it feeds is down. This usually occurs when there is not enough space to queue material between the resources or the resource is actually physically connected to the drum.
5. Poor Throughput rate product mixes. As explained earlier, a key to profit maximization is to make and sell products that produce the most Throughput per time unit on the drum. By making products with a lower Throughput rate, we squander the ability to generate additional cash. There are obvious caveats here as the market may require a company to make a full line of products (each with potentially different rates of Throughput) in order to win any business. (See [Chapter 13](#).)

Other critical factors that affect stability and thus should be measured are the amount of non-constraint overloading and the number of late releases. While TOC expects occasional overloads at non-constraints from time to time, it is important to be able to measure the amount of overload that has and is occurring. If it rises above a threshold (specific to the environment) in the aggregate and at the individual resource area, then the system's stability (and ultimately reliability) will be jeopardized as conflicting

priorities and expedites rise. A late release is work that is released to the production floor after the scheduled released time based on the rope length tied to a drum or shipping schedule. Late releases exacerbate the non-constraint overloads to which we previously referred.

These measures are necessary to encourage localities to use good buffer management and roadrunner techniques (effective subordination) in order to ensure that work is available to the drum at the scheduled time and that drum utilization is protected. Additionally, it encourages problem solving and improvement initiatives in order to protect and bolster uptime on the drums and potential Throughput rates as well as communication to Sales and Management about those Throughput rates.

Metric 3: Speed/Velocity

The objective of this metric is to encourage areas to pass work on as quickly as possible. The time frame in which a system can respond is often a key factor in winning business and effectively managing capital requirements. The iconic basketball coach John Wooden often told his players, “Be quick, but don’t hurry.” Localities must be encouraged to perform work with maximum speed and minimal or no sacrifices to reliability, stability, and quality. If accomplished, it means that the buffer positions that these localities feed can be reduced or the system can be more responsive to potential demand. This metric often takes the form of something called cycle time. *Cycle time measures the time that released material spends within an area rather than the standard machine or labor process time.* By measuring cycle time, a locality is encouraged to enforce the roadrunner rules,¹⁰ encourage movement in time rather than batch, limit WIP inventory, limit early releases, and practice good BM. Conventional metrics like Lead Time, Cycle Time, and

Stock/Inventory turns can also be used to reinforce this objective.

Metric 4: Strategic Contribution

The objective of this metric is to encourage areas to maximize the Throughput rate and Throughput volume according to the relevant factors of the environment and system. As mentioned previously, the relevant factors have everything to do with the defined constraints or leverage points. *Key specific measures of Strategic Contribution will include measuring against a targeted Throughput rate as well as total Throughput.* This metric is designed to encourage all areas to be proactive about participating in the generation of the company's opportunities (e.g., innovative ways to in-source or outsource based on market conditions as well as adding free products) or find ways to increase the Throughput rate (e.g., product or tooling innovation) by creating a feedback loop to measure how well we executed against our plan to exploit the constraint. This is simply variance analysis with a TOC twist and has four components: constraint rate variance (time), product mix, volume variance, and Throughput dollar variance.

The *Throughput dollar variance* is the *budgeted selling price minus budgeted variable costs for a product family compared to the actual selling price and actual variable costs for the product family at the budgeted constraint volume.*

The *product mix volume variance* is the *budgeted volume of the product family versus the actual volume of the product family sold at the standard Throughput dollar rate for the product family.*

The *constraint rate variance* is the *standard constraint rate (planned time on the constraint) for the product family versus the actual constraint time spent on the product at*

budgeted Throughput dollars (selling price minus variable costs per product).

Variance analysis is not proactive; it is a forensic look at the past so we can understand how we used our constraint and judge our exploitation performance.

Remember, the constraint is the primary area where we are measuring utilization and is discussed under “Metric 2: Stability.” While exploitation/utilization of the constraint begins in scheduling, its execution is ensured through BM by identifying effective actions on the shop floor. Having visual loading graphs that clearly show unused/overloaded capacity at the constraint is a proactive tool. The objective is to take actions to sell or make the decision to store the capacity in strategic stock buffers, offload if necessary, or have sales make the call on prioritizing the constraints, workload and communicating changes with the customer.

Metric 5: Local Operating Expense

The objective of this measure is to encourage areas to maximize the local metrics with a minimal or controlled spend. It essentially seeks to measure the amount of money that an area spends in order to convert raw material into Throughput. A local area should be judged against a targeted OE to Throughput generation ratio, which is defined by the relevant range of the TOC economic model demonstrated in [Fig. 14-4b](#). The TOC break-even model is always governed by the impact or lack of impact on the constraint. These local OEs include things like labor, freight, outside processing, contracted or temporary labor, and expedite-related expenses like overtime and premium freight. With regard to this metric, localities will have to balance the level of local OEs with their other critical metrics identified previously. Certainly, a locality should be encouraged to improve flow and velocity with no additional

expenditures. Along the same lines, a locality should not be penalized for increases in OE if they improve the ratio (this actually works in concert with its strategic contribution). This hints at a concept called variable budgeting. Variable budgeting allows areas to increase expenditures based on exceeding their relevant range of volume.

Metric 6: Local Improvement/Waste

The objective of this metric is to point out and prioritize lost opportunities. Specifically, it is measuring a locality's ability to identify an opportunity to move the other local and global metrics in the right direction with minimal or no conflict. Essentially, are we asking the right questions and getting the right answers? One very important aspect of determining this is with reason codes. As described previously in the section on Reliability, a buffer system must collect reasons when work orders enter in the red, late, and early zones.

The required transactional data from the execution of BM can be used to direct improvement efforts including Lean and Six Sigma events and capital application. By forcing reason codes when transactions (receipts) are made in certain key zones (late, expedite, and early) of the buffer and comparing them over time, we can get an amazingly clear picture of how to direct improvement efforts. [Figure 14-8a](#) shows an example of what this picture can look like.

[Figure 14-8b](#) shows some typical types of reason codes for work orders received in the late, red, and early zone receipt and what some potential recommended actions might be.

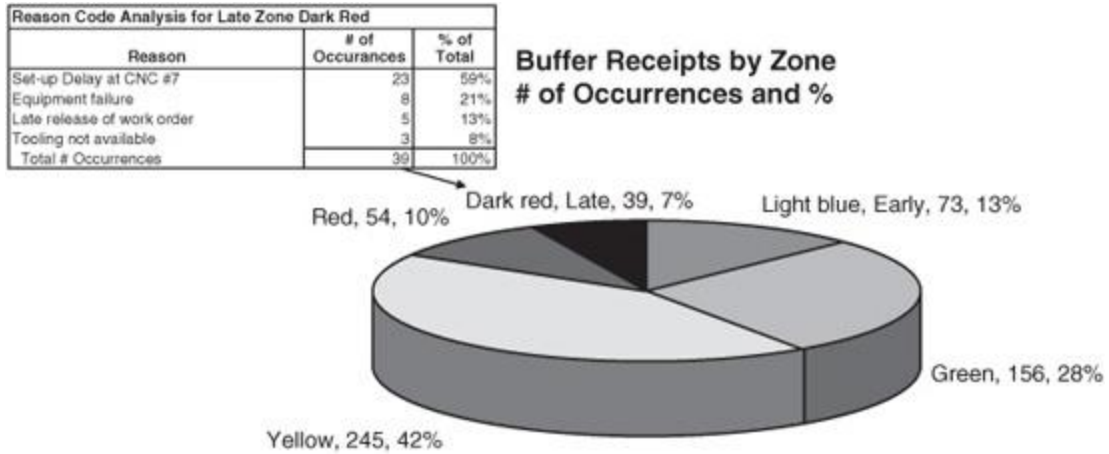


FIGURE 14-8a Reason code analysis.

Zone Receipt	# of Occurrences	Reason	Recommended Action
LATE	23	Setup delay at CNC-Lathe 7	Setup Reduction at CNC-Lathe 7
RED	27	CNC-Mill 18 down	Preventative Maintenance at Mill 18
EARLY	52	Released on time—beat standard	Clean up standards on named routings – evaluate changes on buffers

FIGURE 14-8b Zone receipts with reason codes.

General Measurement	Objective	Specific Examples in Operations
Reliability	Measure execution compliance to plan/schedule	Service levels (on time, fill rates, buffer performance) TOC Model Accuracy (inventory, routing, and standards accuracy)
Stability	Pass on as little variation as possible	Drum Schedule Stability Drum Utilization Non-Constraint Overloading (released and unreleased work) Late Releases
Speed/Velocity	Pass on work as fast as possible	WIP Turns Stock Turns Lead Time Cycle Time
Strategic Contribution	Maximize Throughput rate and Throughput volume according to relevant factors	Targeted Throughput Rate Total Throughput
Local Operating Expense	Maximize above measures with minimal spend	Operating Expense to Throughput Generation Ratio (short-term and long-term relevant range manipulation)
Local Improvement/Waste (Opportunity \$)	Point out and prioritize lost opportunities	Reason Code Collection and Analysis Over Top of Green Stock (exceed the stock buffer limit) Expedite Related Expenses Out of Stock with Demand

FIGURE 14-9 A summary of the six general local measurements.

It is important to note that we are directing improvement to the tails of the buffer zones to capture the largest outliers causing disruption and variation. Too early, our cycle time is too long and excess WIP inventory exists. Too late and we incur overtime and premium freight as well as jeopardize our market promise reliability. By focusing investment/improvement on the tails, we can eliminate the sources of the variation and safely shrink our buffers (time, stock, and capacity).

[Figure 14-9](#) shows a summary of the six general TOC local measurements and their respective objectives as well as some specific examples in Operations.

Feedback and Accountability Systems

Now that we have laid down the foundation for a system of global and local metrics that should point the organization and its localities in the right direction with minimal conflicts, there is still one critical piece of the puzzle left to discuss. The *APICS Dictionary* (Blackstone, 2008, 97) defines a *performance measurement system* as, “(a) system for collecting, measuring, and comparing a measure to a standard for a specific criterion for an operation, item, good, service, business, etc. A performance measurement system consists of a criterion, a standard, and a measure” (© APICS 2008, used by permission, all rights reserved.)

The performance standard can be the accepted, targeted, or expected value.

What is not evident in this definition is the need for steady feedback of system performance and regular adjustment to the actions needed to achieve the standard. The only certainty facing most organizations is that conditions do not

stay the same. For example, a shift in the constraint due to changing market conditions or exploitation efforts will result in the need to modify activity dramatically. Without an effective feedback mechanism contained within the measurements, people tend to drive toward the target without recognition that the conditions of the measurement have changed. In other words, it will result in actions that, even though are believed to be the right thing for improving ROI, can actually be hurting the company. The problem is that yesterday the same actions may have been absolutely the right thing to do. Everyone throughout the organization must understand that the feedback mechanism drives measurement away from being “fixed.” The problem that people can have in understanding this is that the target can stay constant, but the means to achieve it, and therefore the measurement, may change. BM clearly makes this connection for people. Although on-time delivery to the buffer is the target, very different actions are needed every day dependent on the real-time state of all of the buffers (time, stock, and capacity). Decisions on where to flex labor or where to direct maintenance, quality, or engineering efforts may change according to the status of the buffer.

Though an effective operational planning and control solution is a prerequisite to a proper measurement system, the operational system will fail to properly execute or sustain without an effective way to provide feedback on the current system status as well as help to synchronize decisions and actions.

So, How Is the Operational System Performing?

Two very different, and potentially conflicting, approaches to performance measurement exist for answering this question, although both are important. The first approach is using a *performance standard*. A *set goal or benchmark is*

provided, which the employees collectively strive to meet over the course of some finite amount of time. For example, “decrease inventory by 30 percent company-wide in the next six months.” The second approach is to obtain the everyday pulse with an *exceptions feedback mechanism, analyzing the information and deciding if and what action needs to be taken to correct the situation or cause of the exception.* The problem is that while both have a place in organizations, they are easily confused. Any growth opportunities will be minimized when they are up against the fixed performance target for employees’ attention, unless the connection between the two is clear—which often is not the case, especially in larger organizations.

A performance standard will generally create a status quo that an individual being measured will be satisfied with attaining, often ignoring the other factors that are necessary to the optimization of the whole company ROI. This is not the behavior that the organization truly wants and needs because the standard is usually a subset of one of the five tactical objectives of ROI. In other words, the measure will drive organizational conflict (as discussed in the beginning of the chapter). Therefore, once the feedback system directs attention to the source of the problem, the key is to identify, define, and resolve the system conflict. (See conflict resolution in [Chapter 24](#).)

The local metric must be clear, aligned with the global target, deconflicted as much as possible with other measures, and must remain that way. Remember, despite having properly selected productive measurements to begin with if there is not an effective feedback and accountability system providing the current reality and relevance of all local measures to the goal, the system will often be desynchronized and conflicted.

Focusing on Improvement

In contrast to the fixed target, a feedback system does not have an end point but provides continual monitoring of flow to determine exceptions. The regions in BM used to monitor flow are set such that one can respond to an exception and react quickly enough to maintain the desired flow. Additionally, by conducting an analysis of these exceptions and identifying and eliminating the causes of exceptions, a process of ongoing improvement is achieved. Identifying, analyzing, learning, and improving the system is the only way to reach the goal of making the most money now and in the future. By definition, an effective feedback system considers all of the tactical objectives that determine ROI simultaneously rather than any single performance standard because understanding their interdependent nature is a necessity in the feedback system. Companies that understand this thrive on TOC and continue to grow regardless of the economic circumstances in which they are operating. Those TOC implementers who do not understand the significance of BM in providing a process of ongoing improvement will commonly see any improvement stagnate and decline after experiencing initial “brilliant” results and will ultimately end up discarding TOC.

If a manager gains visibility—through BM or any other mechanism—to a potential problem in advance of it affecting performance, this is a great sign that the system is working. Do not mistake this sign with an absence of problems. Companies and the humans who work in them have no shortages of problems. We want those problems to surface when they affect the company performance so that they can be clarified, understood, and resolved. Each problem seen and understood is an opportunity for improvement. Too many individuals of a “fixed” mindset view the presentation of the problem as an indication that the system is *not* working. To accept that identification of potential problems is vital to the measurement and

execution system working effectively is to accept full responsibility and accountability. This thinking is not entirely comfortable for everyone. Without top management understanding and owning this view of the system, there is very little hope that the rest of the organization's management will be able to adopt the "right" mindset.

What Should a Good Measurement System Achieve?

A measure is simply a reading at any point in time of the state of the system relative to the standard the system was directed to execute. It is not used to reward or punish individuals. Cross-functional and interdependent parts of a supply chain can affect the same data but for very different reasons. A fair and productive performance metric will focus and coordinate the efforts of a team, department, process, etc., but a real-time exception feedback is needed to identify exceptions and their causes. Buffer metrics and rules are used to create an early warning system and provide a feedback loop to alert people when and how to act together to get the production system back on plan (to meet market demand). Strategic buffers and BM are used to identify and focus on local improvements most needed for organizational improvement and that have the highest ROI.

It is impossible to separate the measures from the system in TOC because the system *is* the decision-making tool and buffer status reporting is simply the feedback loop on the health of the system. The key to a successful measurement in a BM system is to generate the "hunger" to identify and learn from the problems. The performance standard (when properly aligned) will resolve itself naturally and should not require constant attention from the individuals executing the plan.

Physicist Niels Bohr defined an expert as "a person who has made all the mistakes that can be made in a very

narrow field.” Mistakes, problems, and disruptions in logistical systems should never be regarded as negative unless we do not resolve them, learn from them, and ultimately get better. These are opportunities. Managers should strive to be experts in what they manage. If things are running smoothly, the productive manager is going to push the system to ensure that the buffer is “stressed.” This mindset will undoubtedly result in very short-term negative blips in buffer performance, but will ultimately trend upward and create the learning and thinking organization necessary for ongoing improvement.

The Key Feedback Information

The TOC information system has five necessary components:

1. Constraint and shipping buffer reporting that includes reason code analysis and constraint rate analysis over time.
2. Replenishment/*Actively Synchronized Replenishment* (ASR) buffer reporting that analyzes frequency of zone penetration and records stockouts, stockouts with demand, expedites, and the resulting impact on the shop floor.
3. Pricing indifference modeling (the comparative rate that different products generate cash over the constraint) based on Throughput constraint rates.
4. Strategic market analysis that focuses on both tactical short-term market exploitation (utilizing “free product” capacity) and the mid- and long-range strategic market offers.

5. *Throughput Accounting* (TA) financial statements.¹¹

In today's globally competitive environment, new decision-making tools are required to monitor, measure, and improve the business. A TOC information system is designed to plan, execute, and focus/prioritize improvement. Buffers provide the cushion at strategic points in the production system and BM provides real-time exception reporting on the status of the system. These buffers are visible across the organization and tie local actions to satisfying market demand.

How is the status of the buffers used to ensure sustained improvement? Five questions (5 Q's)¹² must be asked concerning all buffers (time,¹³ stock, and capacity) in the system:

1. What is the condition of the orders? Are they on time or late? Are the Replenishment/ASR buffers healthy?
2. If they are late, is the trend getting worse or better?
3. If it's getting worse, what is the recovery plan?
4. Is the recovery plan effective?
5. What preventive measures are in place to keep the root issue from recurring?

Employing an effective feedback and accountability system requires answering these questions daily, weekly, monthly, and quarterly at different levels of the organization. In TOC, the point is to get an operational system in place quickly that can deal with variation, and a feedback system that can begin the process of execution, feedback, and ongoing improvement.

Generally, there are some key points of measurement and feedback that are important to maintain a *real-time* feedback system. In constraints management there will be

relatively few control points—constraints and buffers—across an entire supply chain that can provide all of the information you need to judge the health of the entire chain and direct attention to places of need and opportunities for improvement.

Remember, all improvements are a change, but not all changes are improvements. If removing variation, waste, setups, etc., does not affect the rate at which Throughput can be generated or speed to market, do not be fooled that the company has made an improvement.

A Problem Is Identified, Now What?

Traditionally, problems are defined as those things in our organization that are not up to par from the fixed target perspective (i.e., “our on-time delivery is too low,” “too much expediting,” etc.). In TOC, these types of issues are referred to as symptoms or undesirable effects (UDEs) given that there is something more fundamental, the core problem, causing these symptoms.

The real problem is something that blocks the symptoms from being permanently addressed—the conflict. If it were as simple as taking an action to combat the symptom, management likely would have long ago resolved the issue. The fact that a symptom still plagues the organization is evidence that there is an equally important pressure—likely another system target or measure being jeopardized—that prevents a sufficient and longstanding solution to prevail. As stated earlier, a good measurement should drive the quest for increased ROI. Given that the tactical objectives for increasing ROI create conflict, resolving conflict must be at the heart of the improvement discussion.

Any time individuals in the organization are not synchronized around the right action to take, then by definition, they are wasting the capacity of their resources.

Conflict over direction and the right action inhibit exploitation of and subordination to the constraint and likely casts an organization-wide doubt in leadership's ability. Once gaining the necessary visibility through the proper execution system, the conflict cloud must become part of the team's toolset to effectively resolve any day-to-day or deeper organizational dilemmas as well as aligning all key members in a common direction.

A common dilemma that many organizations face on a daily basis is the conflict between Sales (top side) and Operations (bottom side) shown in [Fig. 14-10](#).

Sales versus Operations Conflict Cloud

A major customer comes to the salesperson and wants to place a very big order, but it is in a much shorter lead time than currently quoted. The salesperson's job (increase sales) is to book orders for the company to bring in revenue. To do this, he must commit to this short lead time (which disrupts the schedule). On the other hand, Operations has to respond to making orders for all customers (effectively flow product). Operations is continually pressured to expedite parts through the system and has numerous other orders already late or near late. Flow is constantly being disrupted by changes in the schedule (therefore, Operations wants to maintain the schedule). If you have Sales and Operations in your system, you've probably experienced this or some derivative of this conflict. So what's the right answer? Generally, people in Operations will fight for stability in the schedule and people in Sales will fight equally hard for the additional sales opportunity. Does this conflict have anything to do with a fixed performance measure within these departments (i.e., commissions, efficiencies, etc.)?

A market constraint dictates one set of assumptions and will direct the solution in one direction, and an Operations constraint dictates another set of assumptions and will direct the solution in another direction. Without this knowledge, there is no way to properly resolve this cloud. Even if you can resolve the conflict, with fixed metrics driving Sales (sales revenues, sales quotas, commissions) and opposing metrics driving Operations (DDP or overtime, for example) departments independently, there is no way to resolve this cloud to everyone's satisfaction. However, with visibility of the constraint's current load, Sales can be an active participant in both managing sales to exploit the constraint capacity as well as prioritizing the use of scarce capacity when the constraint is overloaded (market spike) in the short-run. A good BM system dramatically reduces the conflict in the organization by giving everyone the same view of the state of the logistical system and tying all of their measures/actions to the global metrics (ROI or RACE).

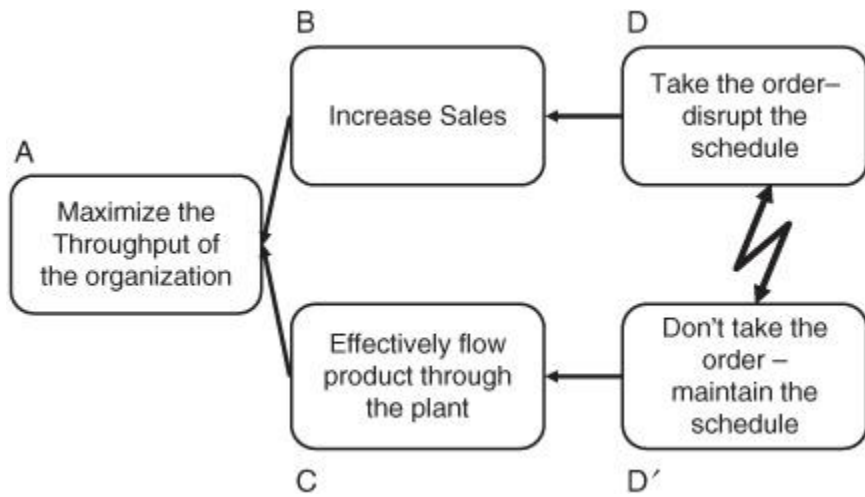


FIGURE 14-10 Sales versus Operations conflict cloud.

Should We Ever Be Satisfied?

We believe that an organization is either growing or dying and therefore should never be satisfied with maintaining the status quo (however healthy the current organization is). The problem is that targets, standards, and metrics are often assumed to have an end point—a state of “being achieved”—and therefore do not promote ongoing improvement. A reference environment that we can use to clarify the problem comes from a professor examining young students’ responses to different forms of expectation and measurement of their academic performance.

Dr. Carol Dweck, a Stanford psychology professor, has been researching the subject of learning and motivation for years. In a series of experiments, Dweck tested the effects of praise and acceptance of achievement (promoting fixed intelligence) versus the effects of praise of hard work and encouraging an interest in tackling adversity (promoting growth). Her thesis concluded that students who hold a “fixed” theory are mainly concerned with how smart they are—they prefer tasks they can already do well and avoid ones on which they may make mistakes and not *look* smart. In contrast, people who believe in an “expandable” or “growth” theory of intelligence want to challenge themselves to increase their abilities, even if they fail at first.

“I also became very interested in coping with setbacks,” she said. . . . “(being) so concerned about not slipping, not failing” (Trei, 2007).

We find a very similar situation concerning fixed organizational targets and standards set as departmental goals. Meeting the target that was set becomes the only concern of the department managers, killing any motivation for ongoing improvement beyond sustainment.

Individuals who excel in education, sports, and industry instantly relate to Dweck’s findings. Tiger Woods, for example, coming off arguably one of the greatest seasons in golf history, made the decision that it was the right time to

completely reconstruct his golf swing. Was this foolish? Not if you understand that Woods' motivation is not driven by the recognition of being the best in the world or a fear of falling from that status. He simply is obsessed with seeking perfection in a game where such a goal is unattainable.

A Case Study

Let's explore another example. A company is vertically integrated and owns its supply chain from raw material through assembly of finished product to be delivered to the dealer or directly to the end user. Purchased parts from outside vendors feed different levels of the *bill of materials* (BOM), but the rest of the process is internal although managed at different plants in different geographical regions. The BOM for major end items is deep (10 to 20 + layers) and, to most people, this would be considered a very complex environment to manage.

Obviously, organizations like this one that are of any size will be broken into manageable pieces to be directed and operated by different individuals. How are meaningful measures provided to the parts of the whole so that they act as one? This company, facing the heights of complexity and delivery challenges, had to take a first step that most of the team would have argued to be the exact opposite of minimizing this complexity—tear down the walls that separated the organization into different business units. It was a necessary condition to any alignment of action and improvement to strip out the systems and metrics that encouraged the whole organization to be viewed as the sum of its parts. This local viewpoint drove organizational conflict over the use of its shared resources (i.e., capacity, inventory, etc.)

Once the artificial segmentation of the organization was removed, the pool of capacity was available to be directed

to the highest need and Throughput opportunity for the company as a whole. Because *customer tolerance time* (CTT) was less than some of the very long lead-time parts were, it was necessary to design and implement a global ASR system immediately followed by DBR (see [Chapter 12](#) on ASR).

Highly visible buffers at only the control points of the organization let all management see the real-time status of the performance of the entire company. No matter how large and complex, the simplicity of TOC allows relatively few points of data collection to provide the relevant information for focusing all decision making. BM and the five questions become the primary day-to-day measurement of the health of the system. Most importantly, the entire organization's measurements are synchronized from local to global through measuring the resources of each feeding link to its buffer. Every cycle time reduction allows for a reduction in the stock buffers supporting its feeding links. For an in-depth understanding and a case demonstrating the dramatic effect this can have on a supply chain, see [Chapter 12](#) on ASR.

Given the size and complexity of this organization, this company designed a central planning function to oversee the trends of the strategic buffers. This allowed for leveraging capacity system-wide, an objective feedback system for upper management, recommendations for improvement initiatives (with the supporting evidence), and an accountability loop to ensure follow-through to the five questions. It was vital that this team also become wellversed in conflict clouds and the tactical thinking processes (Clouds, Negative Branch Reservations [NBRs], and prerequisite trees [PRTs])¹⁴ to mentor and assist other managers in proper alignment for continuous flow.

[Figure 14-11](#) shows the metric and feedback/accountability system at this company.

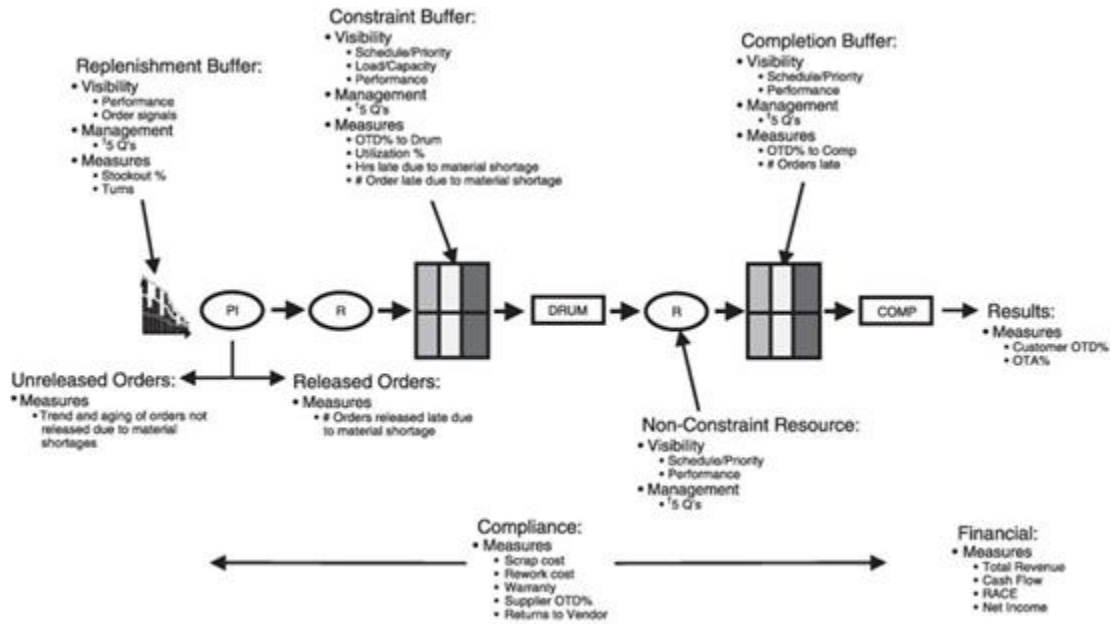
With little more than the tools to provide proper visibility and focus, this company, starting in 2004, was able to exploit a market opportunity to grow from approximately \$260 million to \$1.2 billion and increase RACE from under 5 percent to over 22 percent ([Figure 14-11b](#)). Equally important was that properly focused growth also allowed for better positioning to weather the storm of a global economic downturn in late 2008.

These results were presented at the 2008 Constraints Management User Conference and the 2008 TOCICO Conference, Las Vegas, Nevada (Dan Eckerman, LTI President, *A Vertically Integrated Supply Chain Case*). For additional case study information on this company, see [Chapter 12](#).

Summary

What are the key steps all companies should take to achieve an effective measurement system?

1. Design and implement the proper operational solution given their relevant business factors. In many cases, this means ASR and DBR. Without a clear understanding of the organizational leverage points and their interactions with each other, there is no hope to align actions with a measurement system. Even if the constraint is in the market, it is important to get your house in order first to generate and enable market offers.



¹ 's Q's refer to the five questions discussed earlier in the chapter under the section: The Key Feedback Information (p. 394).

FIGURE 14-11a Metric and feedback/accountability system.

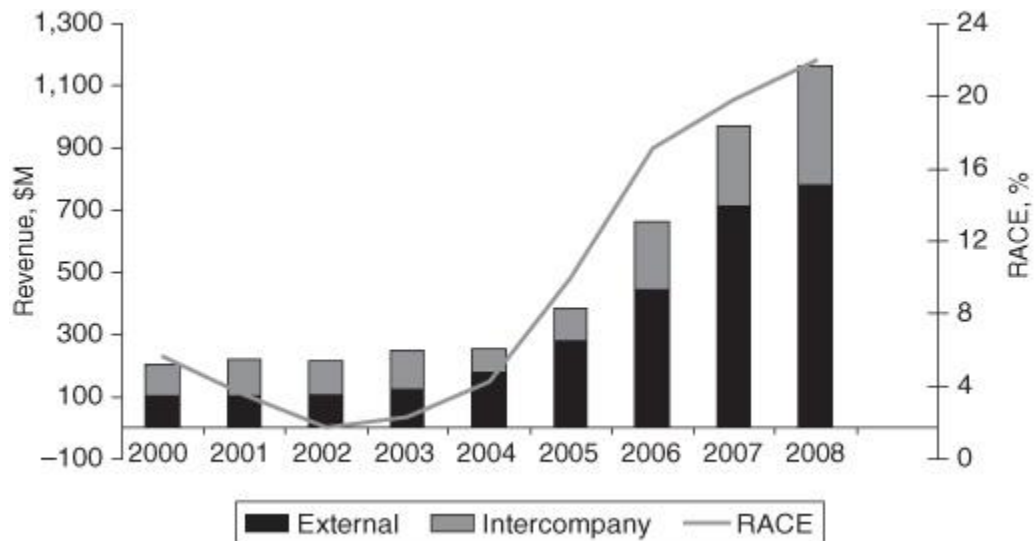


FIGURE 14-11b Growth in revenue and RACE percent.

2. Implement a set of simple and coordinated global and local metrics based on the form of the above solution.

3. Establish highly visible buffers, whether through the use of manual or software mechanisms. These tools are critical to having real-time status of the leverage points of the system (organization or supply chain). With visibility to these buffers, a company can utilize the five questions to ensure proper trending and improvement.
4. Use the TOC tactical thinking process tools, particularly the conflict cloud. Design and implement an information system to provide the reports mentioned earlier. This information is required to align the actions; the conflict cloud provides the framework for organizing and analyzing that information so that everyone understands clearly what actions need to be taken.

The most important thing to remember for all individuals playing a role in a TOC system is that this is a *thinking and evolving system*, not “fire and forget.” Fixed metrics will often point toward a single direction that, regardless of the need of the system, will continue to motivate efforts independently toward its achievement. A truly effective TOC measurement system will point everyone in the direction that will have the greatest return which, by definition, is a growth model. The BM feedback system will provide the relevant information to make day-to-day decisions in line with the organizational ROI measure. As changes occur, people must think, adjust, and adapt to achieve the greatest potential of the organization.

References

Blackstone, J. H. 2008. *APICS Dictionary*. 12th ed. Alexandria, VA: APICS.

- Goldratt, E. M. 1990. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Crotonon-Hudson, NY: North River Press.
- Hornigren, C. T., Sundem, G. L., and Selto, F. H. 1993. *Introduction to Management Accounting*. 9th ed. New York: Prentice Hall.
- Horovitz, B. 2009. "CEO profile: Campbell exec nears 'extraordinary' goal," December 26, *USA Today*.
- Johnson, H. T. and Kaplan, R. S. 1987. *Relevance Lost: The Rise and Fall of Managerial Accounting*. Boston: Harvard Business School Press.
- Mabin, V. J. and Balderstone, S. J. 2000. *The World of the Theory of Constraints: A Review of the International Literature*. Boca Raton, FL: St. Lucie Press.
- Smith, D. 2000. *The Measurement Nightmare, How the Theory of Constraints Can Resolve Conflicting Strategies, Policies and Measures*. Boca Raton, FL: St. Lucie Press.
- Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://tocico.i4a.com/i4a/pages/index.cfm?pageID=3331>
- Trei, L. 2007. "New study yields instructive results on how mindset affects learning," Stanford Report. Stanford, CA: Stanford University. Available at: <http://news-service.stanford.edu/news/2007/february7/dweck-020707.html>.

About the Authors

Debra Smith is a partner with Constraints Management Group, LLC, an international partnership committed to assisting organizations achieve breakthrough results and sustainable, ongoing improvements using the Thought Process (TP) tools and application solutions offered through the Theory of Constraints (TOC).

Ms. Smith has extensive experience in public accounting, financial management in manufacturing companies, teaching at the university level, and consulting in TOC. Debra began working with Dr. Eli Goldratt in 1990 when she was an Associate Professor of Accounting at the University of Puget Sound. She is responsible for original research in the field of TOC applications in manufacturing environments and has created numerous courses and workshops integrating TOC and traditional manufacturing measurement and scheduling processes.

Her research has focused on understanding changes necessary in measurements, accounting and information systems to support continuous improvement processes in manufacturing. She is coauthor of "The Theory of Constraints and Its Implications for Management Accounting," an independent research study of TOC funded by the Institute of Management Accounting, and she is the author of a book entitled *The Measurement Nightmare, How the Theory of Constraints Can Resolve Conflicting Strategies, Policies and Measures* (St. Lucie Press, 2000).

Prior to teaching, Ms. Smith worked in public accounting as a CPA for Deloitte & Touche and spent nine years in publicly traded manufacturing firms, both as a Division Controller and Vice President of Finance and Operations. She is internationally recognized as an authority on management accounting and is a noted speaker on TOC.

In 2001, Ms. Smith was elected to the founding Board of Directors of the Theory and served for five years on the Theory of Constraints International Certification Organization (TOCICO), a certification organization founded

by Dr. Eli Goldratt. Ms. Smith is certified by TOCICO in all applications of TOC (Operations Management, Distribution Management, Project Management, Finance and Measures, TOC Thinking Processes, and Holistic Management.), since 2003.

Jeff Herman has dedicated the past fifteen years of his work to the development and practical application of strategic and tactical Thinking Processes with an emphasis on resolving organizational conflict.

Mr. Herman began working with the TOC Thinking Processes in 1994 and received extensive formal training in TOC from the Avraham Y. Goldratt Institute Academy. Following his education in the Academy, Mr. Herman went on to become a Jonah's Jonah and worked as a Regional Director and Product Specialist in the United Kingdom for the Goldratt Institute for two years. He opened a new territory for TOC implementations with Dr. Eli Goldratt in the Baltic States in 1997.

Mr. Herman returned to the United States in 1998 and now resides in Eau Claire, WI. He is currently a partner of Constraints Management Group, LLC (CMG)—a leading international consulting enterprise that specializes in the application of TOC—and is Practice Leader of Strategic Thinking Processes. Since the late 1990s, Mr. Herman and his partners at CMG have been at the forefront of developing and articulating the concepts behind Actively Synchronized Replenishment (ASR) as well as building ASR and Drum-Buffer-Rope (DBR) compliant technology. He has guided hundreds of executives and managers through the application of strategic and tactical thinking processes within a diverse scope of organizations and industries across 16 countries.

Mr. Herman is a certified expert by the Theory of Constraints International Certification Organization in the TOC fields of Operations Management, Distribution

Management, Project Management, Finance and Measures, TOC Thinking Processes, and Holistic Management.

Mr. Herman and Ms. Smith are currently coauthoring a revision of *The Measurement Nightmare*, due to be released in the fall of 2009.

CHAPTER 15

Continuous Improvement and Auditing

Dr. Alan Barnard

Introduction

The Goal—Achieving Continuous or Ongoing Improvement

Fundamental to the success and viability of any organization is a realization (by the management team) that improvement is not a once-off event and that continuous or ongoing improvement requires continuous change. Unfortunately, not all changes result in improvement and continuous changes can jeopardize stability. Ensuring that every significant change results in an improvement (in both performance and stability) for the organization as a whole is one of the most significant challenges faced by the management of any organization. It requires a reliable focusing mechanism to differentiate between all the *many* parts and processes that *can* be improved from those *few* that *must* be improved (to achieve more organization goal units now and in the future).

Dr. Eli Goldratt (1986) became one of the continuous improvement pioneers in the modern era with his book, *The Goal*. Its subtitle hints that the real goal for organizations is not just to make more money now and into the future, but

simply to ensure the organization is on a “*Process of Ongoing Improvement*,” or POOGI, to achieve sustainable growth and stability. Achieving POOGI in any organization not only requires a reliable focusing mechanism (to identify where and what to change and when and what not to change), but also a holistic decision-support mechanism (to judge the system-wide or global impact of changes). Then, a fast and reliable feedback mechanism is needed for auditing progress/compliance or for identifying other important system performance gaps or variations. Even more importantly, it requires a different mindset and thinking about improvement at all levels in the organization to systematically identify and challenge the policies, measurements, behaviors, and underlying assumptions that limit the current organizational performance.

Copyright © 2010 by Dr. Alan Barnard.

In the introduction to *The Goal* (1986), Goldratt describes such a process:

Finally, and most importantly, I wanted to show that we can all be outstanding scientists. The secret of being a good scientist, I believe, lies not in our brain power. We have enough. We simply need to look at reality and think logically and precisely about what we see. The key ingredient is to have the courage to face inconsistencies between what we see and deduce and the way things are done. This challenging of basic assumptions is essential to breakthroughs. Almost everyone who has worked in a plant is at least uneasy about the use of cost accounting efficiencies to control our actions. Yet few have challenged this sacred cow directly. Progress on understanding requires that we challenge basic assumptions about how the world is and why it is that way. If we can better understand the world and the

principles that govern it, I suspect all our lives will be better.¹

One of the major “inconsistencies” relating to the topic of continuous improvement and auditing is why, especially considering the advances and discoveries over the past 100 years in continuous improvement and auditing of organizations and the intense competitive pressures, so many of the changes made in organizations are not sustainable. And why do most changes “fail”—either not resulting in any measurable improvement in organizational goal units or even causing decay in performance to the extent that organizations themselves frequently fail.

Purpose and Organization of This Chapter

This chapter aims to provide a framework for designing a continuous improvement and auditing process within organizations from a *Theory of Constraints* (TOC) perspective and to share some of the important new TOC developments in this field since *The Goal* was first published in 1984. The chapter starts with the definition of key concepts and a brief historical perspective on this subject. It then provides an overview of the current gap, extent, and consequences (vicious cycle) related to traditional continuous improvement and auditing methods and mistakes (why change). We then examine, the underlying conflicts and assumptions that need to be challenged, (what to change), the solution criteria and direction and details of a solution to break these conflicts and prevent new undesirable effects (to what to change), and finally how to overcome the typical implementation obstacles (how to cause the change) to implementing such a TOC-based continuous improvement and auditing solution.

Key Concepts and Definitions

Continuous improvement (CI) is defined simply as *the continual improvement (in organizational or system goal units) over time*. CI can also refer to the continual improvement of subsystems, processes, or products or services provided by an organization, but with the warning that unless these “local improvements” can or will contribute to improving the organization as a whole, they cannot be called improvements but rather “local optima.” In fact, the Japanese word Kaizen, made famous by Masaaki Imai’s book (1986), *Kaizen: The Key to Japan’s Competitive Success*, is frequently used today as a synonym for CI because the translation of “kai” (change) and “zen” (good) literally means “good change” (improvement for the system as a whole). In the context of this chapter, “continuous” is used to refer to all types of ongoing improvement rather than as a way to differentiate small marginal (low-leverage) improvements from large step-change (sometimes defined as discontinuous or high-leverage) improvements.

Continuous improvement process (CIP) is by definition a *closed-loop cycle of sequential steps designed to bring about continual improvement through a process of discovery, application, review, and corrective action*. The Shewhart cycle (Plan-Do-Review-Act), Six Sigma’s DMAIC (Define-Measure-Analyze-Improve-Control), and TOC’s *Five Focusing Steps* (5FS) are among the best known.

Change impact is classified into three types, with *Type 1* referring to a change that results in a *measurable improvement*, *Type 2* referring to a change that *did not result in a measurable improvement or decay* (within the “noise”), and *Type 3* referring to a change that resulted in a *measurable decay in performance* of an organization as a whole or a specific process output.

Auditing is defined as *an ongoing process of review of an organization, its process, projects, products, services, or*

subsystem's performance and compliance against standards or expectations. In the words of Winston Churchill, "However beautiful the strategy, you should occasionally look at the results." Auditing is an important part of CI in any organization as it provides a practical feedback mechanism for stakeholders with the objective to reduce the time to detect and time to correct performance gaps, variations, or noncompliance. It is in this more general context that the terms "audit" and "auditing" would be used in this chapter rather than the more common use where "audit" refers only to internal or external financial auditing. As part of a CIP, there are typically three types of audits that are done. *Compliance auditing is the organization doing what it should be doing (and not doing what it should not be doing).* Performance auditing is the organization performing as well as it is expected to perform. *Potential auditing is the expectation that the organization do (much) better.*

A Historical Perspective—Standing on the Shoulders of Giants

The desire and capability to continuously improve our lives and understandings of the systems with which we interact have played a critical part in the evolution of our species. But it was not until the development of the "scientific method"—initially formulated by Aristotle around 350 BC and improved upon through significant contributions by the likes of Ibn al-Haytham (965-1040), Roger Bacon (1214-1294), Francis Bacon (1561-1626), Galileo Galilei (1564-1642), René Descartes (1596-1650), Isaac Newton (1643-1727), John Stuart Mill (1806-1873), and more recently Karl Popper (1902-1994)—that there was a systematic way to challenge and continuously improve our assumptions, knowledge, and methods to analyze, improve, manage, and predict causes and effects within a specific type of system. The *scientific method* is simply defined as *a systematic or*

iterative method in which a problem or objective is identified, relevant data is gathered, a hypothesis is formulated, and the hypothesis is empirically tested (e.g., through validation of effect-cause-effect predictions) and then improved upon after review of the experimental test results. The scientific method allows scientists to test theory and methods with experimentation and allows them to use insights gained from experimentation to develop new or improved theories or methods.

Some of the most important discoveries to advance our knowledge and methods of CI and auditing of organizational performance have been made by the likes of Taylor, Gilbert, Ford, Shewhart, Deming, Juran, Ohno, and Goldratt, who knowingly or unknowingly simply applied the scientific method to the science of analyzing, improving, managing, and predicting the performance of organizations.

Many of these discoveries capitalized on the importance of reducing overall process *time delays* and later, the importance of reducing *quality defects, process variation, lost time on capacity constraints* and *overproduction* to improve the overall performance of the system. Benjamin Franklin's famous quote, which he shared with a young tradesman in 1748, "Remember that time is money," specifically referred to the opportunity cost of wasting time on something that could be done faster, with less defects or that should not have been done at all. Simply stated, *slow processes* (or ones that contain defects or variation) *are expensive processes* (George, 2002).

These discoveries resulted in powerful CI methods such as the *Toyota Production System (TPS)*, *Lean*, *Total Quality Management (TQM)*, *Six Sigma*, *Business Process Reengineering (BPR)*, and *TOC*—each with a large reference bank of success stories and "best-practices" that could provide a baseline for auditing (e.g., the ISO 9000 family of standards for auditing TQM systems).

But with such a powerful and tested toolkit of CI and auditing methods, one would expect that the adoption rate of these tools would be very high and that the majority of those who really tried to implement these methods and tools would achieve major jumps in performance compared to past results.

Why Change?

Introduction

Despite the impressive reference bank of successes and the powerful insights of today's mainstream continuous improvement methods, they all seem to struggle with achieving higher levels of adoption, with sustaining and expanding on initial improvements, and probably most importantly, with finding ways to reduce the significant percentage of failures and wasted scarce resources due to these failures.

This section provides an overview of the analysis for answering the question "Why change?" (the conventional way) by reviewing the typical improvement gaps within many organizations (and individuals), starting with a common improvement challenge and then a literature review to quantify the extent, consequences, and vicious cycle related to the high failure rate of most "improvement" initiatives within private and public sector organizations today.

The Improvement Gap and Challenges

There are many differences between types of organizations and within organizations from the public and private sectors.

However, all goal-orientated organizations (and individuals) have two characteristics in common:

1. They are complex systems (many parts and many interdependencies between the parts) that make them difficult to analyze, improve, manage, and predict the impact of change.
2. There is continuous pressure to achieve more (goal units) with less (resources) in less time resulting in conflicts such as “do what is good for the short term vs. do what is good for the long term” and “do what is good for one part vs. do what is good for other parts (the system).

[Figure 15-1](#) shows an example of this pressure and challenge to improve resulting from a large growing gap between stakeholder expectations (the “red” curve) and actual performance (the “green” curve).

For private sector organizations, this challenge manifests in the continuous pressure to close the gap between actual and expected short- and long-term returns for shareholders. For public sector organizations, the challenge manifests itself in the ongoing pressure to close the large and frequently growing gap between the deteriorating levels of service delivery and infrastructure and a growing demand for such services in the areas of health, safety, education, energy, and telecommunications—especially in the developing countries around the world. For individuals, the challenge manifests itself in the difficulty to maintain a balance within the various aspects of our lives—some struggle with gaps in their self-confidence, others with gaps within their health, some with gaps in their relationships, and others with gaps in financial security.

Organizations and individuals also share three types of responses to such pressure to change due to current and

likely future performance gaps and unacceptably high variations that can cause system instability:

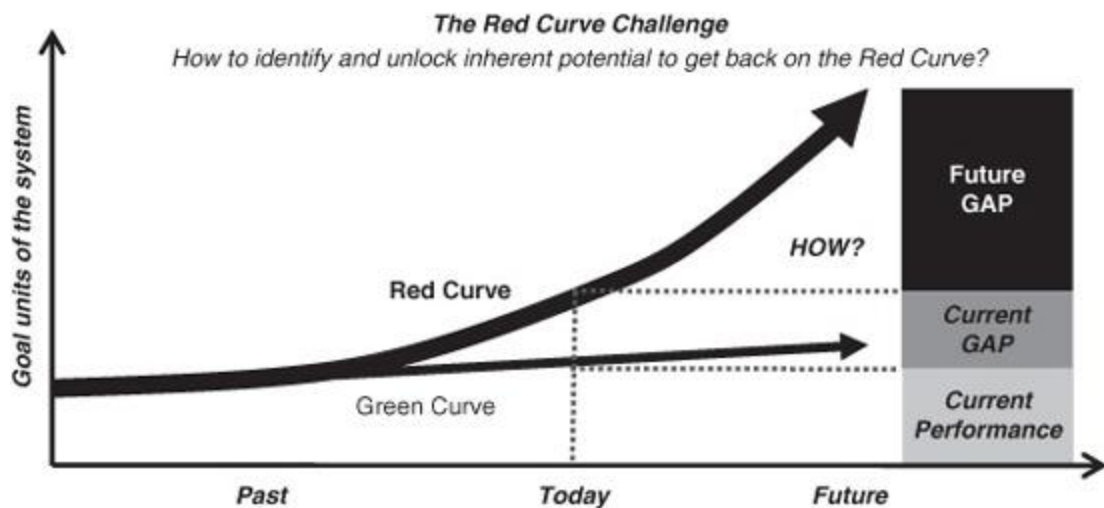


FIGURE 15-1 The red curve challenge. (© E. M. Goldratt used by permission, all rights reserved. Source: Modified from Goldratt, 1999).

1. Don't change (to prevent decay or at least to prevent wasting resources).
2. Make many small- or low-leverage and low-risk changes (to maintain stability).
3. Make few large- or high-leverage and possibly high-risk changes (to achieve growth).

[Figure 15-2](#) shows the uncertainties and the related conflict that determines which of the three responses will be the most likely for a specific stakeholder. When organizations (and individuals) are faced with the reality that their performance is no longer improving at the required or desired rate or have unacceptable high variation, they face the risk of performance decay if they don't change (the uncertainty of not changing). At the same time, if they decide to change but "play it safe" by targeting

many small incremental improvements, they will probably risk not meeting their growth objective, while if they decide to target the few large step-change improvements, they risk instability and even decay, which could threaten the survival of their organization (the uncertainty of changing).

These uncertainties put all stakeholders who feel or are held responsible for the performance of the system into the conflict on the right-hand side of [Fig. 15-2](#). In order to achieve ongoing success, stakeholders feel they must meet the required or desired growth objectives. In order to meet these growth objectives (to reduce gap or variation), they feel pressure to change. At the same time, to achieve ongoing success, stakeholders also feel they must ensure that the requirements for stability (and survival) are never compromised, which contributes to the pressure not to change or at least not to initiate any step-changes that could jeopardize stability and even survival.

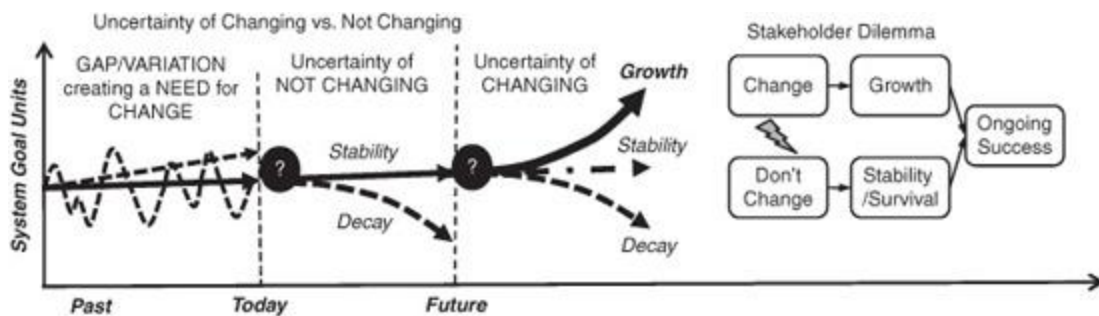


FIGURE 15-2 The uncertainty and dilemma related to the improvement challenge.

The Types of Management Mistakes When under Pressure to Change

The design of a continuous improvement and auditing system (to create a learning organization) should start with classification of the types of mistakes made that can block continuous improvement. There are two types of mistakes²

(Ackoff, 2006): *Errors of commission, doing something that should not be done or not doing the right thing properly; and errors of omission, not doing something that should have been done.* Ackoff warned that we learn little from doing things right or even from doing the right thing at the right time. Most learning comes from doing the wrong thing or doing something wrong. However, in order to learn from such mistakes, they must first be detected, their cause or source must be identified, and a solution must be developed to prevent such mistakes in the future. Unfortunately, in most organizations mistakes (especially errors of omission) are hidden, sometimes even from those who made them.

But what percentage of the changes made by management result in measureable and sustainable improvements that meet the expectations of all stakeholders (Type 1 impact) versus what percentage of changes fail to meet measurable objectives (Type 2) or cause decay in performance (Type 3)?

The Extent and Consequences of the Failure Rate of Change

The extent of failure rates of different types of improvement or change initiatives together with the extent of organizational failures can provide a good indication of the consequences of errors of omission and commission.

The Failure Rate of Improvement/Change Initiatives

A representative sample of research studies and surveys (listed in [Table 15-1](#)) shows that regardless of the type of change initiatives, between 50 and 80 percent of these initiatives fail to meet their original objectives, are stopped before completion, or sometimes even cause the organization's performance to decay. The only study that formally reported that no failures or disappointments were

reported was conducted by Mabin and Balderstone (1999) involving implementations of TOC at 100 companies.

Analysis of the studies reporting high failure rates shows that the vast majority of the changes are reported to fall into the second category of change impact—where there is neither a direct measurable benefit nor decline. Of course, the “cost” in these cases is not only the wasted costs or investments incurred (without benefit), but also the wasted opportunity costs of not applying scarce resources (especially “management” time—the real constraint in most organizations) in changes that would have improved the system performance. *This also does not mention the impact of such a high failure rate on people’s reduced motivation and expectations for future changes.* Considering such a high percentage failure rate of change initiatives, what is the failure rate of companies and organizations?

Failure Rates of Companies

When it comes to companies, research studies show that failures are also statistically much more likely than successes. Since the advent of the modern corporation, over 10 percent of all companies in the United States (the largest and most successful economy in the history of the world) fail every single year; 22 percent of the top 100 companies at any given time drop from the elite rankings in the next decade; and 50 percent of globally successful companies go extinct within the lifetime of a modern human (Ormerod, 2006, 13). A study by the U.S. Census Bureau (www.sba.gov/advo/research/data.html) showed that 25 percent of new businesses started in 1992 failed within the first year and by year 10, the failure rate was 70 percent.

Change Initiative	Number (N)/ Failure Rate %	Study Objective	Top 3 Reasons for Failure
TQM A. D. Little, 1992	64 (ADL) –80 (ATK)	Survey of 500 companies by Arthur D. Little, survey by A. T. Kearney of 100 UK companies	Lack of top management support, resistance to change
Six Sigma Angel & Pritchard, 2008 R. Farrelly, 2008	60	"What went wrong with Six Sigma? A look into Six Sigma's 60 percent failure rate," July	Resistance to change and lack of top management support
	> 50	The top reasons Six Sigma Projects fail: Results from a survey of Six Sigma Projects in 114 companies, Ross Farrelly, AOQ Six Sigma Conference, Melbourne, August 2008	Partial implementation, project not linked to ROI, poor management support
Lean R. G. Kallage, 2006	Well over 50	"Lean implementation failures: Why they happen, and how to avoid them," Richard G. Kallage, July 11, 2006	Lack of top management support/poor business case, resistance to change, and poor deployment
Balanced Score Card G. DeBusk, 2006	70	"Does the balanced scorecard improve performance?" Study published in <i>Management Accounting Quarterly</i> , Fall 2006, Gerald K. DeBusk	Not available

Business Process Reengineering Dr. Malhotra, 1998	55–70	"Business process redesign: An overview," BRINT Institute	Resistance to change (associated with downsizing)
Organizational Transformation J. Kotter, 2009	70	"Leading change: Why transformation efforts fail, survey of 100 companys' transformation efforts," John P. Kotter, HBR, March 2009	Resistance to change, lack of urgency, lack of support from top management
Outsourcing Initiatives Gartner	63	A Gartner Group survey of 180 clients in 1995 on failure rate of outsourced IT arrangements	Not available
New Product Launches Linton, Matysiak, and Wilkes, 1995	70–80	Research study by Linton, Matysiak, and Wilkes, Inc. on failure rate of new product launches for the retail grocery industry (reviewed 1935 new products introduced by the top 20 U.S. food companies)	Lack of good R&D (market resistance), poor or under-resourced execution

MRP/ERP Robbins-Gioia Survey, 2001	51	ERP System Implementation Success Survey	Data inaccuracy, lack of top mgt support, resistance to change
Other IT Projects Chaos Report, 1994 vs. 2004	Cancelled /Fail: 18, Challenged: 53	IT Project Failures—Chaos Reports by The Standish Group. Results from 1994 versus 2004 study show no significant improvement	Lack of user or top management support, resistance to change, unclear user requirements
TOC	15 not sustained 15 not started	Research study by Realization Inc. on "Critical chain projects: Successes, failures, and lessons learned," presented at TOCICO Conference in 2005	Lack of buy-in into the necessary three TOC rule changes, and not establishing "how-to" mechanics
	0	A review of Goldratt's Theory of Constraints (TOC)—Lessons from the international literature, Mabin and Balderstone, 1999	"In the survey of over 100 cases, no failures or disappointing results were reported"

TABLE 15-1 High Failure Rate for Various Change Initiatives and IT Projects

Whenever we see such large failure rates, it is quite likely that there is some vicious cycle at work where actions taken with the intent to correct a situation have the opposite effect. The next section provides insights into the vicious cycles seen in many organizations that are not improving at the desired rate or within those that no longer exist (i.e., the ones that experienced catastrophic failures).

The Vicious Cycle Related to the High Failure Rate of Change

Many of the studies reviewed not only quantified the extent of the high failure rate of change, but also analyzed the most likely causes and consequences of the high failure rate. The consequences of the high failure rate mentioned in most studies are no surprise—higher resistance to change for future initiatives and lower expectations for the likely impact of future initiatives. What might be surprising (unless readers have experienced these themselves) is that there is also a remarkable consistency in the findings of the different studies (regardless of type of change initiative) on the major

reported causes. Most of the studies list the two main causes of the high failure rate as “Resistance to change” (especially by middle managers) and “Lack of active support or under-resourcing by top managers.” This is reportedly caused by the project team’s relatively low expectations of the likely benefit of the proposed change (i.e., what you can’t quantify, you can’t justify the allocation of scarce resources). But these two factors are the same as what was identified as the consequences of the high failure rate.

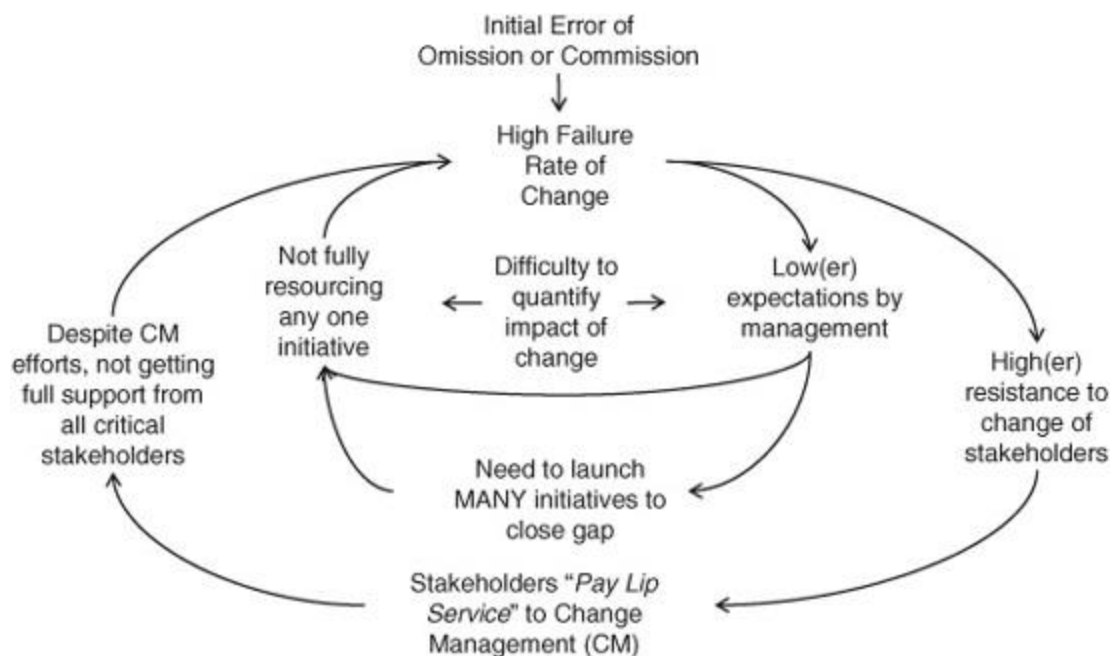


FIGURE 15-3 Vicious cycle related to high failure rate of change initiatives.

When a specific behavior is both a consequence and a cause, it means the system is likely to be stuck in a *vicious cycle* (Senge, 1990, 80-83) such as shown in [Fig. 15-3](#). The higher the failure rate, the higher the resistance and the lower the expectations of stakeholders. Moreover, the higher the resistance, and the lower the expectations, the more likely those necessary changes will be blocked or the necessary changes will not receive the full support and

resources needed to make them a success, which again increases the probability for failure. Over time, a vicious cycle such as this stabilizes and soon those trapped within the cycle conclude that a response of “it will never work” is a safer response than embracing new changes or that simply, considering the complexity and uncertainties within their system, this (high failure rate) is probably the best they can do.

This fear related to the high failure rate of changes can also explain why changes that focus on local cost, waste, or process variation reduction (low-leverage changes) are more likely to be supported because they are perceived to be lower risk and more certain. High-leverage changes that focus on “changing the rules” are less likely to be supported because they are considered to be high risk and less certain.

Summary of Why Change?

In summary, the literature review on the failure rate of change initiatives found that regardless of the type of change (with the exception reported in the study on TOC projects), whether it was implemented within the private or public sectors, and/or how many inspiring successes have been reported related to that type of change, change/improvement initiatives are far more likely to fail than to succeed. Studies that have been repeated, such as the Chaos Report on IT Project Failure rates, also show that despite the significant insights gained and widely reported as to the consequences and causes of these failures during previous studies, the failure rate has not changed measurably, making it a much safer option for stakeholders to resist change or pay “lip service” during the launch but walk away saying “it will never work.”

Why Change?—The Gap		
Typically, 70 percent of all change initiatives fail and some important changes are never implemented.		
No.	Undesirable Effect	Why Is It Bad?
1	Many causes/constraints outside your control	Can result in inertia and complacency
2	Organizations are very complex	Difficult to identify root causes/where to focus
3	Uncertainty on impact of change/not change	Difficult to justify whether to change or not
4	Too many change initiatives going on at once	Conflicting priorities and fights for resources
5	High resistance to change to new initiatives	Results in conflicts and “paying lip service”
6	Lack of top management support	Many projects not fully resourced or started
7	Stakeholder expectations not always clear	High rework/change not meeting expectations
8	Slow or no feedback loops on changes made	Long time to detect and time to correct mistakes

TABLE 15-2 Summary of “Why Change?”

[Table 15-2](#) provides a summary of “Why Change?” in the format used in a TOC analysis, which includes a clear problem statement (the gap) and the *undesirable effects* (UDEs)—the effects that stakeholders would complain about that makes it difficult to close the gap (solve the problem).

However, what should be changed to eliminate or reduce these UDEs in the way managers identify, plan, execute, and audit CI and other change initiatives?

What to Change?

Introduction

The fact that executives and managers keep trying new strategic or process improvement and other change initiatives despite their abysmal rate of failure is, as per Samuel Jackson’s famous quote, “Like second marriages, a triumph of hope over experience.” On the other hand, it may indicate just how much pressure top managers face to improve the performance of their organizations. The large

failure rate of improvement methods triggers the classic innovator's dilemma (Christensen, 1997)—most innovations fail, but companies that don't innovate might die.

No wonder there have been calls such as “Innovate or Evaporate” (Tucker, 2002). But why are many necessary changes not implemented or implemented in time (errors of omission) and why does the high failure rate of implementing change persist (errors of commission) despite our evolving understanding of the cause-effects that govern ongoing success or failure of organizations? This is the question we will try to answer in this section.

Finding the Core Conflicts within Continuous Improvement and Auditing

In science, there is general consensus that by “defining a problem precisely, you are halfway to a solution” (Goldratt, 1990, 37). Goldratt proposed a method called the “Evaporating Cloud” (EC; sometimes referred to as a Conflict Cloud or Conflict Diagram) to provide a practical mechanism for “defining a problem more precisely” by verbalizing the unresolved problem as an unresolved conflict in trying to satisfy two different sets of necessary conditions within the same system. By understanding the conditions that create the conflict (underlying erroneous assumptions about the system and behavior of its parts), we can gain insight as to what few changes will be needed to solve the core problem—the few changes that would “evaporate” the core conflict cloud of the system and therefore reduce or even eliminate the performance gap and related UDEs. [Figure 15-4](#) shows an example of the core conflict in deciding which organizational structure to use—the conflict of centralize versus decentralize. In order to achieve success (A—The common objective in the conflict), managers must ensure the organization is *efficient* (B—a necessary condition for success). In order to ensure the

organization is *efficient*, management feels pressure to centralize (D—Assumed prerequisite for satisfying the necessary condition of *be efficient*).” At the same time, to achieve success the organization must be effective (C—another necessary condition for success), which results in pressure to decentralize (D’—Assumed prerequisite for satisfying the necessary condition of *be effective*). However, if they centralize too much, some stakeholders will complain about increased bureaucracy and slower decisions (the negative consequences of jeopardizing the need for *be effective* or “Not C”), which they believe can be corrected by decentralization. However, if you decentralize too much, other stakeholders will complain about increased noncompliance and duplication or waste in common resources (the negative consequences of jeopardizing the need for *be efficient* or “Not B”). This method of using the conflict cloud to *better define the problem* helped in this case to understand that the real problem is the unresolved conflict resulting in oscillation between centralization (to prevent noncompliance and waste) and decentralization (to prevent bureaucracy and slower decisions). This centralize/decentralize conflict and its consequences can be seen in many organizations today and will continue, until this conflict can be broken.

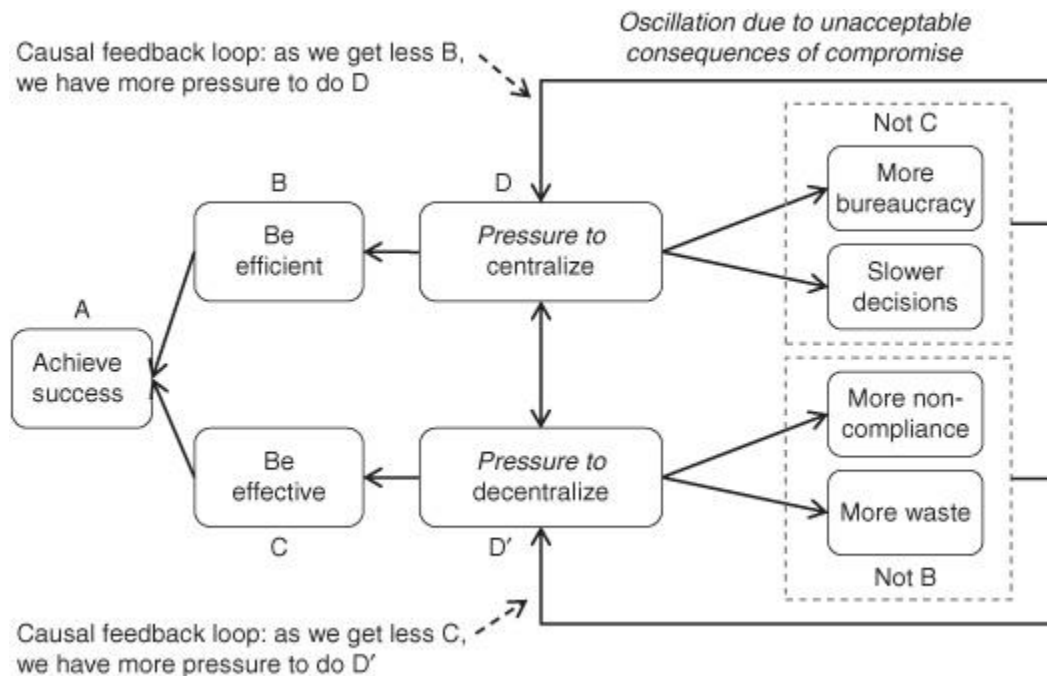


FIGURE 15-4 Example of core conflict within an organizational structure.

However, what are the unresolved core conflict(s) faced by managers related to achieving ongoing growth and stability in their organizations? Step 1 in the process to “define the problem more precisely as an unresolved conflict or set of conflicts” is to identify the UDEs or generic bad decisions related to the errors of omission and commission in CI and auditing:

1. Not changing when you should or changing when you should not—mistakes in deciding on *When to Change*.
2. Implementing the wrong change (e.g., unimportant/nonurgent changes) or not implementing the right change—mistakes in deciding *What to Change*.
3. Implementing the right change in the wrong way (e.g., without full consensus or not fully resourced)—

mistakes in deciding *How to Change*.

4. Not correcting or stopping (a change) as soon as possible when we recognize one of the above three mistakes were made—mistakes in *auditing changes*.

Step 2 simply involves verbalizing the actions/decisions related to each of these UDEs as part of an unresolved conflict. In Box D, we write the action we feel the most pressure to take when dealing with the problem. Box D' represents the (opposite) action that caused the problem. Boxes B and C are the needs each action is trying to satisfy (or the needs that will be jeopardized if the actions in D and D' are taken) and last, Box A is the common objective or goal for that system or subsystem. As an example, if the UDE/problem is a growing performance gap, then D—the action to deal with the problem—is “Change now” to satisfy a need (B) of “Improve Performance/Stop Decay.” The opposite action (D') is “Don't change now” to satisfy a need (C) of “Maintain Stability/Personal Security” and the common objective is “Ongoing Success.”

[Figure 15-5](#) shows the three generic (core) conflicts for *when to change*, *what to change*, and *how to change* (including *when to stop a change*).

Mistakes of omission (*when* or *what not*) and commission (*what* and *how*) are closely linked. Although mistakes of omission can simply be due to ignorance (e.g., when the change needed is unknown or counterintuitive), the main reason people make mistakes of omission is that they fear making mistakes of commission (Ackoff, 2006). From the outside, it frequently appears as if the assumptions on which such fears or claims of “not knowing” are based are not rational. Therefore, to prevent these mistakes, we need to identify what assumptions are ultimately driving the wrong decisions when faced with these conflicts and then find a way to show that these assumptions can and should be challenged.

Finding a Simple and Systematic Way to Break Conflicts

Leonardo da Vinci (1452–1519) said, “All our knowledge (and decisions) has its origins in our perceptions (our assumptions about reality).” The decisions relating to *when* to change (and *when not* to change), *what* to change (and *what not* to change) as well as *how* to change (and *how not* to change), and whether to *stop* or *rework* are influenced by our individual and organizational assumptions or “paradigms.”

With the TOC *Thinking Processes* (TPs), the key to finding any breakthrough solution is to identify, invalidate, and remove one or more of the “erroneous” or limiting assumptions that block us from breaking the conflict (what to *stop* thinking or doing) and to replace it with a “more valid” assumption that will enable achievement of a better win-win (what to *start* thinking or doing). The simplest and frequently most effective and efficient way to find such erroneous assumptions is to focus on the conflict arrows within each of the core conflict clouds (Barnard, 2007) i.e. Why D jeopardize C, why D’ jeopardize B, why D and D’ is in conflict and why there is not another way (E) to satisfy B and C.

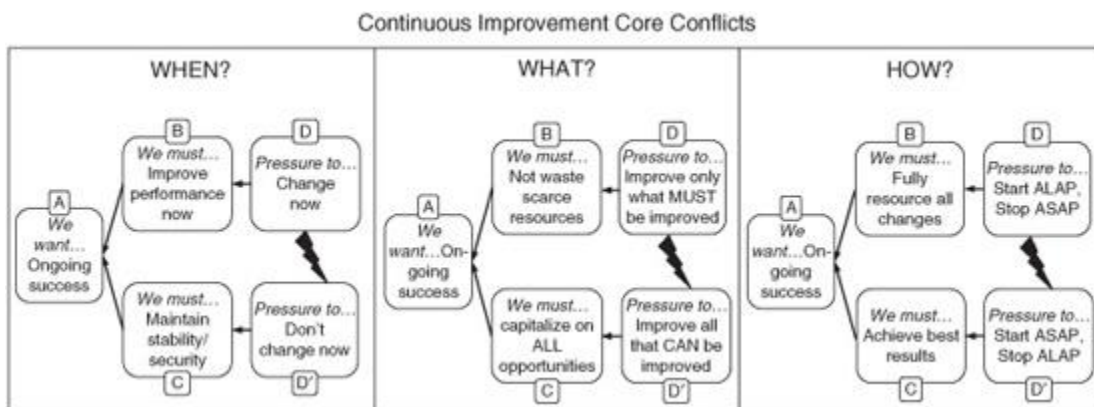


FIGURE 15-5 Core conflicts related to knowing when, what, and how to change.

Challenging Assumptions Related to WHEN (and WHEN NOT) to Change

There will be disagreement on *when* to change and *when not* to change as long as some stakeholders believe it is not possible to change or that they are doing the best they can (due to an assumption of “a constraint that is out-of-my-control”) or that the change is not necessary (due to an assumption of “we still have time”). To break this conflict, we need a reliable way to validate (or invalidate) that the all constraints can be overcome and that we don’t have any more time (without risking serious consequences).

Challenging Assumptions Related to WHAT (and WHAT NOT) to Change

There will be disagreement on *what* to change and *what not* to change as long as some stakeholders believe that more is always better, that every local improvement will result in a global improvement, or that focusing scarce resources on a few high-leverage opportunities is too risky or not fair (i.e., we should capitalize on *all* improvement opportunities). To break this conflict, we need an acknowledgment that management must focus their scarce resources on high-leverage changes, which requires a way to differentiate between all the many parts (of a complex system) that can be improved from the few that must be improved now to get more goal units.

Challenging Assumptions Related to HOW (and HOW NOT) to Change

There will be disagreement on *how* to change and *how not* to change as long as some stakeholders believe that *the earlier we start, the earlier we finish*—an assumption that is true only when we are not bad multitasking. Other related assumptions that will result in this type of conflict is whether

to wait until we get full consensus or can fully resource the initiative, or when some believe that failure is bad and therefore any attempts to audit or stop any changes not making the planned progress should be resisted. To break this conflict, we need a way to validate (or invalidate) that starting new initiatives (that share resources with existing initiatives) sooner will not simply result in both the current and new initiative finishing later (see Chapters 3, 4 and 5 on the effects of bad multitasking). Or further, that not reviewing or not stopping initiatives that are not delivering will be a lose for all stakeholders especially when these initiatives consume scarce resources.

Identifying Limiting versus Enabling Paradigms in Continuous Improvement

We can classify the types of assumptions that need to be challenged by organizations wishing to continuously improve based on the five generic improvement challenges (Barnard, 2007) faced by managers of any form of complex system.³ The assumptions and related beliefs used by managers to decide how to best deal with these five challenges can turn the challenges into either obstacles (that lock-in current performance) or opportunities that will allow managers to see and unlock inherent improvement potential within their organization. The five challenges include:

1. How to deal with *constraints*, especially those considered “out of your control,” when setting targets and expectations for improvement.
2. How to deal with the inherent *complexity* of your organization, especially when deciding where to focus your improvement efforts and scarce resources or

when trying to predict the impact of changes on the organization as a whole.

3. How to deal with strategic and day-to-day policy or resource allocation *conflicts* within your organization between stakeholders from the same or different parts of the system, especially in environments where there is significant distrust.
4. How to deal with the *uncertainty* and potential *risk* when having to decide on which changes are needed, what the impact of these changes will be (on achieving more goal units), when to start these changes (to not trigger bad multitasking or resistance to change), and when to stop a change if there are insufficient resources or it is not delivering the expected benefits (errors of detection and correction).
5. How to deal with “*bad behavior*” of people that has resulted or could result in significant UDEs for the system, especially in cases where the way we deal with such people could have other repercussions (e.g., union strikes, etc.).

We have a choice as to which set of assumptions (paradigms) we use to make decisions related to these five challenges and on what we focus as a result. [Figure 15-6](#) provides a summary of the limiting (traditional/conventional) versus enabling (systems approach/TOC) assumptions or paradigms that govern how a manager will deal with these five challenges and whether the challenge is viewed as a major obstacle or a major opportunity on which to be capitalized.

1. We can assume that constraints are inherent (limiting) or that there is always inherent potential for

improvement—that all constraints can be overcome (enabling).

2. We can assume that the best way to improve complex systems is to break up these systems into simpler parts and improve each part (limiting) or instead, that the best way is to find the inherent simplicity—the constraint in physical flow or the few root causes that explain most UDEs in any system—the leverage points of the system (enabling).
3. We can assume that the best way to deal with conflicts is to compromise or focus on the win for you even if it causes a win-lose (limiting) or we can assume that a win-win is always possible when we collaborate to move from “me versus you” to “us versus the problem” (enabling).
4. We can assume that there is inherent certainty by looking for optima points according to some textbook formula (limiting) or we can assume that since uncertainty is inherent, we should rather find a logical solution and a “good enough” starting point and use feedback to detect and correct cases of “too much” or “too little” (enabling).
5. We can assume that bad choices or bad behavior are made by bad people and that we should get rid of such people (limiting) or that since we believe that people are good, we assume that bad choices or behavior are made by good people with bad assumptions, so we rather find and get rid of the bad assumptions.

Summary of What to Change

There are three generic conflicts managers face in deciding *when, what, and how* to change to achieve ongoing success for their organizations. Within each of these conflicts, there are key assumptions that can and need to be challenged to enable managers to know how to break these conflicts in a win-win way. These “limiting” assumptions when used to make decisions on how best to deal with constraints, complexity, conflicts, uncertainty, and bad choices can result in errors of omission, commission, detection, or correction. A new set of “enabling” assumptions is proposed by TOC (and other systems approaches) that can help prevent these management errors. The famous line of Qui-Gon Jinn of Star Wars: Episode I—The Phantom Menace fame can summarize the enabling assumptions, “Your focus determines your reality.”

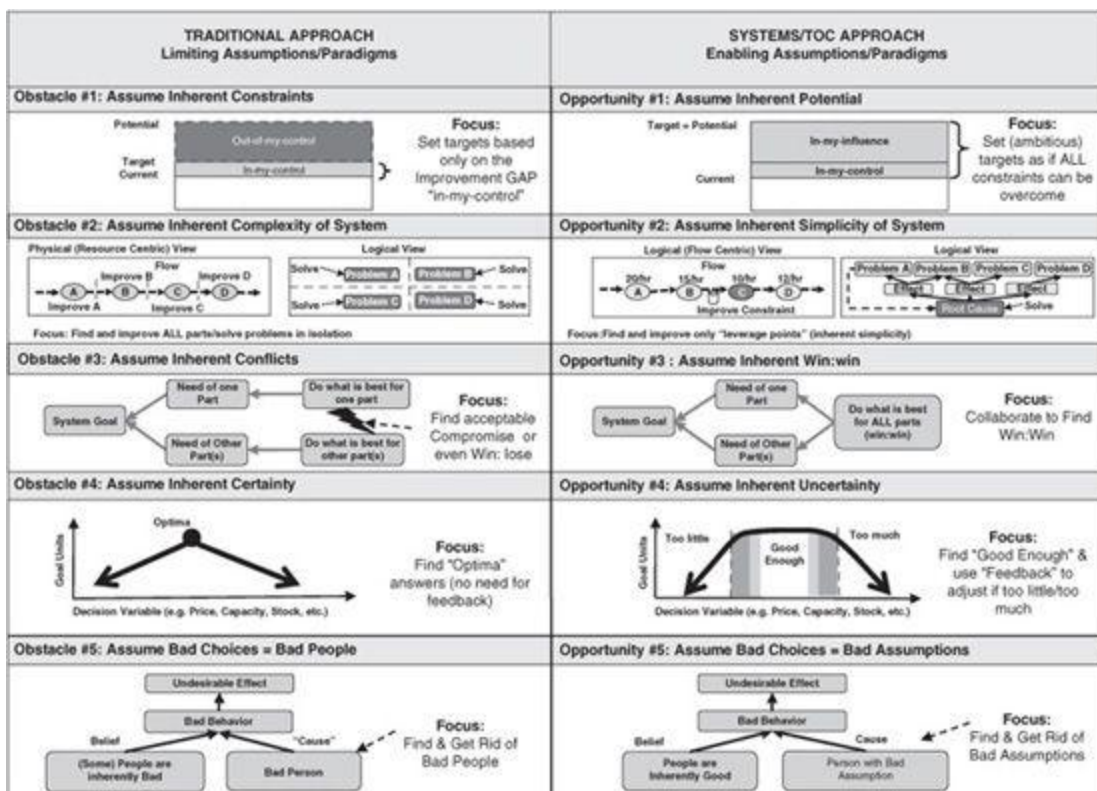


FIGURE 15-6 Limiting versus enabling paradigms to deal with five improvement challenges.

Focus on everything that can be improved and the possible becomes impossible. *Focus* on the few things that must be improved now (to get more goal units), and the impossible becomes possible.

What do we mean by *focus*? *Simply doing what should be done and not doing what should not be done*—the opposite of the mistakes or errors of omission (not doing what should be done) and errors of commission (doing what should not be done), which provides the simple answer to “What to change?”

The next section explores the answer to the question “To what to change?”

To What to Change?

Introduction

To answer “To what to change” in a TOC analysis, we have to answer four questions which we will apply to our analysis on designing a holistic CI and auditing system:

1. What are the criteria we should use to judge a real breakthrough solution?
2. What is the direction of the solution that will break the core conflict and prevent (or at least reduce) the major undesirable effects within the current reality we are trying to improve on?
3. How do we translate the generic solution into a specific solution for various applications?

4. What changes will be needed to prevent the new solution from causing unintended negative consequences (potential UDEs) through either its failure or success?

Criteria to Evaluate a New Solution

In a previous section, we listed the gap and major UDEs that make it difficult to close the gap in both CI and auditing of organizations. Therefore, the criteria of a new solution (the desired effects, or DEs) can simply be stated as the opposite of the gap and UDEs defined in “Why change?” The DEs for a holistic CI and auditing system include:

1. Know where to focus scarce resources for the best result (despite complexity).
2. Provide a way to quantify the likely impact of changes (despite uncertainty).
3. Raise expectations especially with top management to ensure full support (despite obstacles).
4. Each stakeholder can actively contribute to ensure the change will result in a system improvement.
5. No complacency, inertia, or fear of failure.
6. Ongoing alignment/synchronization of contributions toward the goal of the organization.
7. Ensure a higher success rate of change initiatives (rather than 70 percent failure rate, we should target and achieve a 70 percent success rate).
8. Reduce the time to detect and time to correct for wrong assumptions or poor execution.

Direction of Solution to Breaking the Continuous Improvement Conflicts

To ensure that we don't make the common mistakes of omission and commission (within CI) and mistakes of early detection and correction (within auditing), we need a focusing mechanism that helps us identify what should be done (to achieve more goal units) and what should not be done (since it would waste resources or even decay performance).

TOC provides a simple and effective solution to this problem—no wonder that more and more organizations have been looking to TOC to provide them with the focusing mechanism needed to focus all improvements on what is good for the company as a whole (Breyfogle, 2008).

TOC's Five Focusing Steps

Goldratt (1990a, [Chapter 1](#)) defined a simple *Five Focusing Steps* (5FS) process for achieving continuous and step-change improvement that, if followed, would also likely prevent the errors of omission and commission as well as errors of detection and correction. The process is based on the simple premise that an organization can be viewed as a chain and, therefore, any organization's performance is limited by its "weakest link" or system constraint. To improve the organization's performance, management should therefore focus their limited time and resources on finding ways to "strengthen the weakest link." TOC's 5FS process enables an organization to continuously exploit and elevate its inherent potential that can be "unlocked" or "created" through focusing scarce resources on identifying, exploiting, and elevating the performance of its current system constraint. TOC's 5FS are as follows:

Step 1: **Identify** the system constraint (the weakest link).

Step 2: Decide how to **exploit** (not waste the potential of) the system constraint.

Step 3: **Subordinate** everything else to the above decision.

Step 4: **Elevate** the constraint.

Step 5: If, in the previous steps, a constraint has been broken, **go back to Step 1**. WARNING: Do not allow inertia to cause a system constraint.

Frequently, the most difficult step is Step 3—subordinate everything else (all processes, policies, and measurements) to the decision on how to better exploit the system constraint—because it can result in local versus global optima or short-term versus long-term conflicts.

For example, if the constraint is in factory capacity, it might make sense for the factory to produce in larger batches to reduce setup and waste. However, if the constraint moves to the market and the company is starting to lose sales due to the fact that their lead times are too long or their unwillingness to accept smaller orders (both consequences of a policy to manufacture in only large batches), then the factory will face a new conflict. Should they now change the old rules (or not) to subordinate to the requirements of the new system constraint—the market—by producing smaller batches. Unless this conflict is broken, the company will not be able to exploit fully the market potential.

A similar conflict can arise when the company enters Step 4—elevate the system constraint. The company might have a policy in place not to hire any additional people or not to approve any capital expenditures, which, unless this conflict is broken, will block the company from elevating its system constraint and therefore constraining its improvement potential. [Figure 15-7](#) shows a graphical representation of the 5FS and the related exploitation and elevation conflicts organizations might face when realizing they would have to

challenge and possibly change some of the rules to better exploit or elevate the real system constraint.

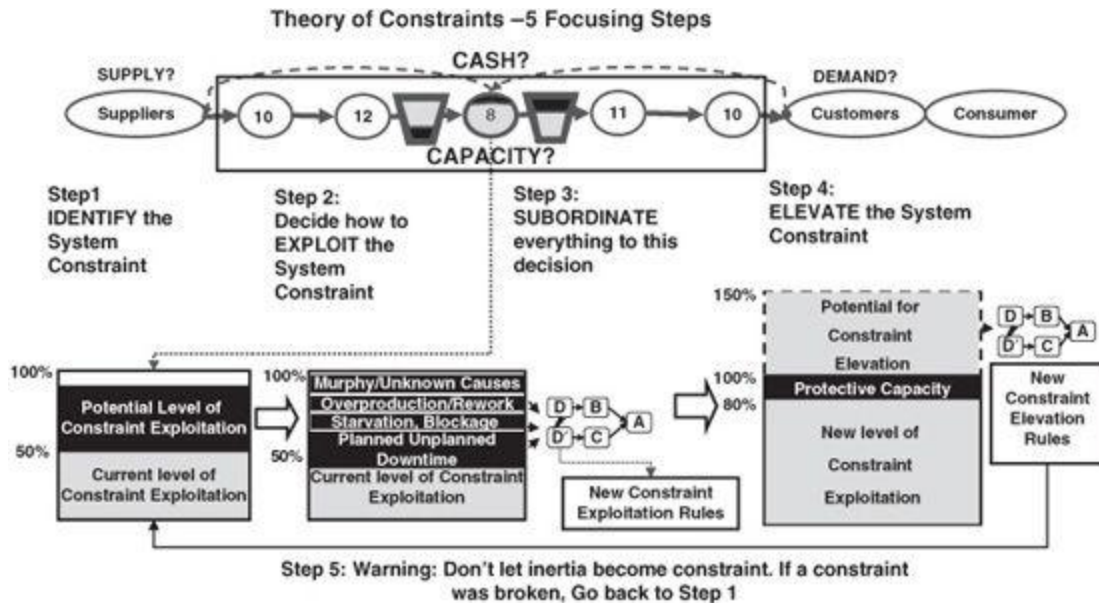


FIGURE 15-7 Graphical representation of the TOC 5FS (Barnard, 2003).

But you may ask, what about the “non-constraints?” Should they not be improved?

The system constraint is the only part of the system for which “more is better” is valid. All other parts (non-constraints) have to maintain their performance at a level of “good enough.” If their performance is below this level (too little), it will compromise the performance of the constraint and therefore needs to be corrected as soon as possible. We call an improvement at a non-constraint “local optima” if the improvement will raise the performance of the non-constraint significantly above the level of “good enough” (too much) and/or cause the performance of the constraint to deteriorate. Therefore, we should never make the mistake of thinking that TOC’s 5FS says non-constraints are not important. Non-constraints are necessary conditions but have to be managed to perform within their required

“threshold” level of good enough—not too much, but also not too little. If they are performing below their threshold, they *must* be improved.

To determine what this threshold level is for non-constraints, TOC uses the concepts of capacity, stock, and time buffers. If the buffer for which a non-constraint is responsible (e.g., Human Resources department to ensure a sufficient pool of skilled craftsmen; Procurement to ensure sufficient stock and acceptable lead times of raw materials and purchased parts) is in the “red,” this indicates that one of two changes in the starting conditions occurred. Either the demand has increased (which means the capacity of the non-constraint might have to be elevated) or the supply performance is less reliable than assumed or is not sufficient to maintain the buffer (which means the non-constraint’s performance must be improved). At the same time, if a buffer is maintained without too much red zone penetration, it means the performance of the non-constraint is good enough should not be improved further until red zone penetration becomes too much (more than 10 percent).

In summary, the 5FS provide a generic process for achieving CI in any organization, and should be the focusing mechanism for all process improvements.

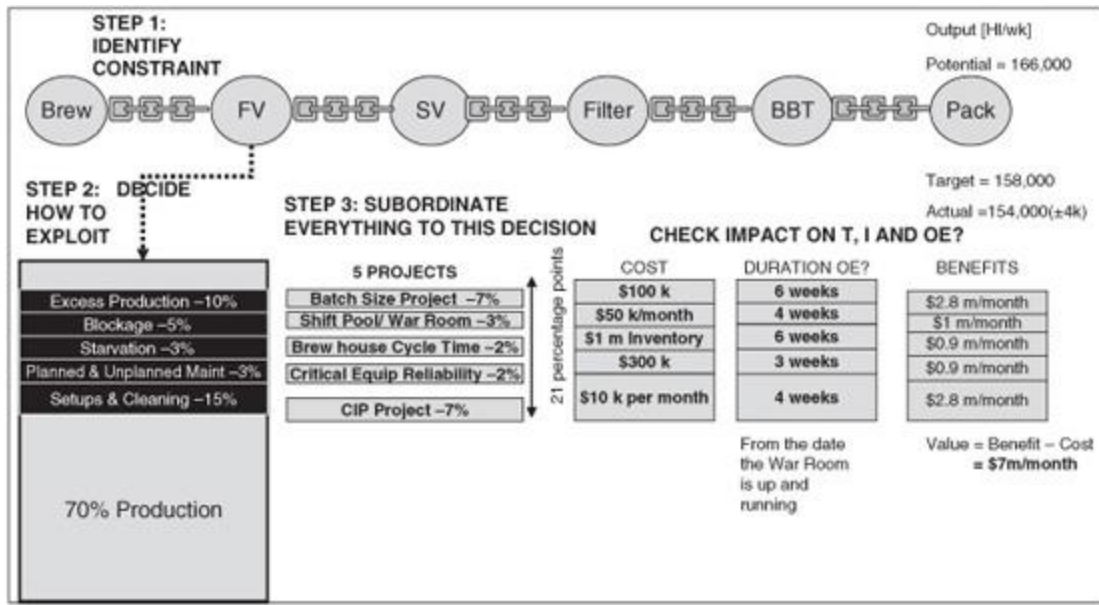


FIGURE 15-8 Applying TOC's 5FS to a brewery.

Applying the 5FS to Achieving CI at a Brewery

[Figure 15-8](#) shows how the 5FS can be used within a TOC-based CI process at a brewery to:

- Identify the flow constraint (fermenting vessels—FV);
- Decide how to better exploit the flow constraint (by reducing capacity lost on setups and cleaning, planned and unplanned maintenance, starvation and blockage, and excess/over-production); and
- Determine the subordination actions—the change initiatives (projects) needed to better exploit the constraint (faster cleaning-in-process or CIP project, critical equipment reliability improvement project, brew house cycle time reduction project, adding a shift pool to buffer against absenteeism, adding a “War Room” to report on the status of buffers every shift, and deciding

on corrective actions and a project to reduce batch sizes to reduce over-production).

Each of these subordination action decisions was the result of breaking a previous subordination conflict that blocked better exploitation of the system constraint. For each project, the impact on Throughput (T), Operating Expenses (OE), and Inventory (I) are calculated as well as the time to implement duration to determine the value unlocked (Value = Benefit - Costs). This one-page summary is referred to as a “Constraint Exploitation Sheet” that can also be used for capturing, communicating, and auditing the impact of the constraint exploitation change initiative.

Applying the 5FS to Develop a Business Strategy—the Viable Vision Process

The most effective way of using the 5FS is when it is applied at the organizational level. In his recent public “Now-and-into-the-future” seminars, Goldratt recommends that in the case of for-profit companies we should start with the assumption that (at the highest level) the constraint for profitable growth is simply management time (bandwidth). But how do we apply the 5FS to management time? To help identify where management should focus (not waste) their scarce time, we should view the market as a strategic constraint and apply the 5FS accordingly.

This means that “Step 2—Decide how to exploit the system constraint” is really deciding what conditions, if satisfied, will get customers to pay more or buy more (i.e., the conditions for building, capitalizing on, and sustaining a decisive competitive edge). As a result, “Step 3—Subordinate everything to the above decision,” requires that process improvement initiatives be focused on only those processes, policies, measurements, or behaviors that block

the company from satisfying the conditions for building, capitalizing on, and sustaining a decisive competitive edge.

In 2005, Goldratt shared the process he uses himself for analyzing companies to determine which few changes are needed now for a company to become an “*ever flourishing company*.”—*a company with exponential performance growth and improved stability*. He called the process the *Viable Vision* (VV) Process and it includes the following six questions:

1. What is the VV growth/improvement target for the company?
2. How much do sales have to increase in order to reach the VV growth target (calculated by determining what price and/or volume increase in sales is possible and subtracting the associated increase in totally variable cost)?
3. Is the existing market large enough to allow the required increase in sales (through either price or volume increase) to be achieved with better exploitation or will it require elevation of the market constraint (with new products into existing markets or existing products into new markets)?
4. How can this increase in sales be accomplished (what conditions, if satisfied, will enable increased price and/or sales volume and what changes are needed to satisfy these conditions)?
5. How much additional capacity, Operating Expenses, and Investment will be required to support this level of sales (by exploiting before elevating)?
6. Can the company (and its suppliers) support the necessary change(s) required to achieve the growth

targets (its management, systems, suppliers, cash, etc.). If not, what additional changes are needed to ensure that non-constraints don't turn into constraints?

This process is aligned with the focusing philosophy that has been applied by many successful CEOs such as Jack Welch, ex-CEO of General Electric. He stated (Pande et al., 2000, 6) that,

The best Six Sigma Projects begin, not inside the business but outside it, focused on answering the question: How can we make the customer more competitive? What is critical to the customer's success? Learning the answer to that question and learning how to provide the solution is the only focus we need.

The VV Process is also in line with the need to understand the cause-effect relationships of which internal changes are needed to increase and sustain higher value to customers and shareholders recommended by Kaplan and Norton (2002, 69) in developing a strategy map on which a balanced scorecard can be based (a cause-effect map showing the relationship between financial, customer, internal and learning and growth perspectives). At the top of the strategy map should be the financial targets (how shareholders will benefit). Below this should be the competitive edges (how customers will benefit and why customers will pay more or buy more). Below this will be the necessary changes in processes and policies to build these competitive edges, and at the bottom of the strategy map should be the necessary enablers to support ongoing improvement and learning.

Change Question	Implementation Process Questions
Q1: Why Change?	Are there any significant gaps or variation in the prime measurements of the system? These gaps and the difficulties we face in closing them are called undesirable effects in TOC.
Q2: What to Change?	How do we differentiate between the many symptoms and the few causes and what really blocks us from addressing these causes (the unresolved conflicts and erroneous assumptions that block us from breaking the conflicts)?
Q3: To What to Change?	What is the direction of the solution that will solve the core problem and resolve the core conflict? In addition, what are the potential negative effects of the new changes (to solve the core problem) and what can be done to prevent these from happening?
Q4: How to Cause the Change?	What are the potential implementation obstacles and what is needed to overcome these? In addition, considering limited resources and interdependencies between the required changes, in what sequence should the changes be implemented?
Q5: How to Measure the Change and Achieve POOGI?	What measurements should be used to determine whether local changes are in fact resulting in a system/global improvement and which measurements and processes should be implemented that will encourage and enable the POOGI?

TABLE 15-3 Simplified Generic Continuous Improvement Process Using TOC's TP

TOC's Thinking Processes

The TPs of TOC were invented to help managers when they get stuck with finding an answer for one or more steps of the 5FS. These TP can also be used in isolation to deal with day-today management challenges,⁴ but are generally used in combination as part of a holistic analysis on an organization or specific subject matter. Goldratt (1990a) originally grouped the analysis/change process into three questions, starting with "What to change?" However, this might create the impression that all stakeholders already agree on the need for change. Since both a literature review (e.g., Kotter, 1990) and field experience show that this is not always a reasonable assumption, we should add "Why change?" as Step 1 if we want to use the TP as a generic analysis and CI and auditing process (Barnard, 2003). In addition, the third (and last) question proposed by Goldratt

was “How to cause the change?”, which does not link back to “Why change?” to create a “closed-loop” framework for a process of ongoing improvement. To close the loop we should add, “How to measure the change and achieve a POOGI?”. The Five Question Change Framework provides a generic CI process as summarized in [Table 15-3](#).

These five questions provide a simple analysis and consensus roadmap for any CI initiative and can generally be presented and applied in a five-day workshops⁵—one day per step as long as all the key stakeholders are present.

Step1 (Day 1) aims to get agreement on the new systems approach (transition from the conventional limiting to the enabling paradigms of TOC). It also includes getting agreement on the answer to “Why change?” for the system and its stakeholders being analyzed by identifying the system’s performance gap, consequences of not closing the gap, and the list of UDEs of each stakeholder that make it difficult for them to contribute to closing the gap.

Step 2 (Day 2) aims to answer and gain consensus around “What to change?” by getting each stakeholder to verbalize their conflict in addressing their UDEs, showing how these are examples of a deeper core conflict, and then identifying possible erroneous assumptions and related policies and measurements that block closing the gap effectively and efficiently.

Step 3 (Day 3) is dedicated to answering “To what to change?”; achieving consensus on which new assumptions and related policies or measurements will break the conflict; remove the UDEs; and close the gap without creating new UDEs.

Step 4 (Day 4) is focused on answering “How to cause the change?” by identifying the possible risks (negative branches and implementation obstacles) and how to prevent or overcome these risks by constructing a sequenced implementation plan.

Step 5 (Day 5) is focused on agreeing how specific contributions will be measured and recognized as well as how stakeholders will know that the gap is really closing, which provides the answer to “How to measure the change and create POOGI?” [Figure 15-9](#) shows a graphical representation of how to use the TOC TP as a generic CI and auditing process.

TOC’s Functional Management Solutions and Their POOGI Mechanisms

The simplified TOC TP analysis roadmap also provides a simple framework for capturing the full TOC analysis and solutions for each of the main functions within organizations: Operations, Finance, Projects, Distribution, Marketing, Sales, Managing People, and Business Strategy. For each of the applications, the *simplified TP roadmap with the five change questions* can be used to communicate a summary of the TOC analysis. This answers the five change questions, including the gaps in prime measurements (how we know improvement in this area is really needed), the typical UDEs that make it difficult to close the gap, the core conflicts, the erroneous assumption and related “old rule” or core problems that should be challenged, the new TOC insight to break the core conflict and related change in policies, measurement, or process (new rules), the steps to implement the change, and finally the POOGI.

[Figures 15-10](#), [15-11](#), and [15-12](#) show the summary of the TOC analysis and CI opportunity and solution for managing Operations, Projects, and Distribution/Supply Chain. Appendix A contains the TOC analysis and CI opportunity and solution for managing Finance, Marketing, Sales, People, and Business Strategy.

These templates can be used as an auditing tool to identify CI opportunities within your organization. If your organization suffers from the performance gap and the

UDEs stated in any of the “Why change?” boxes, it is likely that the associated TOC solution (Answers to “What to change?” “To what to change?” and “How to cause the change?”) can provide a simple and powerful way to unlock inherent potential. The way to measure and achieve ongoing improvement and the mechanisms needed to achieve this are defined in the “How to create the POOGI” boxes.

Details of TOC’s Buffer Management to Focus Ongoing Improvement Efforts

In “Standing on the Shoulders of Giants,” Goldratt (2009)⁶ suggests that the key to the success of Henry Ford and Taiichi Ohno was the fact that they built their management philosophy and planning and execution rules around what Goldratt calls “*the four concepts of supply chains*”:

1. Improving flow (or equivalent lead time) is a primary objective of operations.
2. This primary objective should be translated into a practical mechanism that guides the operation when not to produce to prevent overproduction. (Goldratt showed that Ford limited space; Ohno limited inventory.)

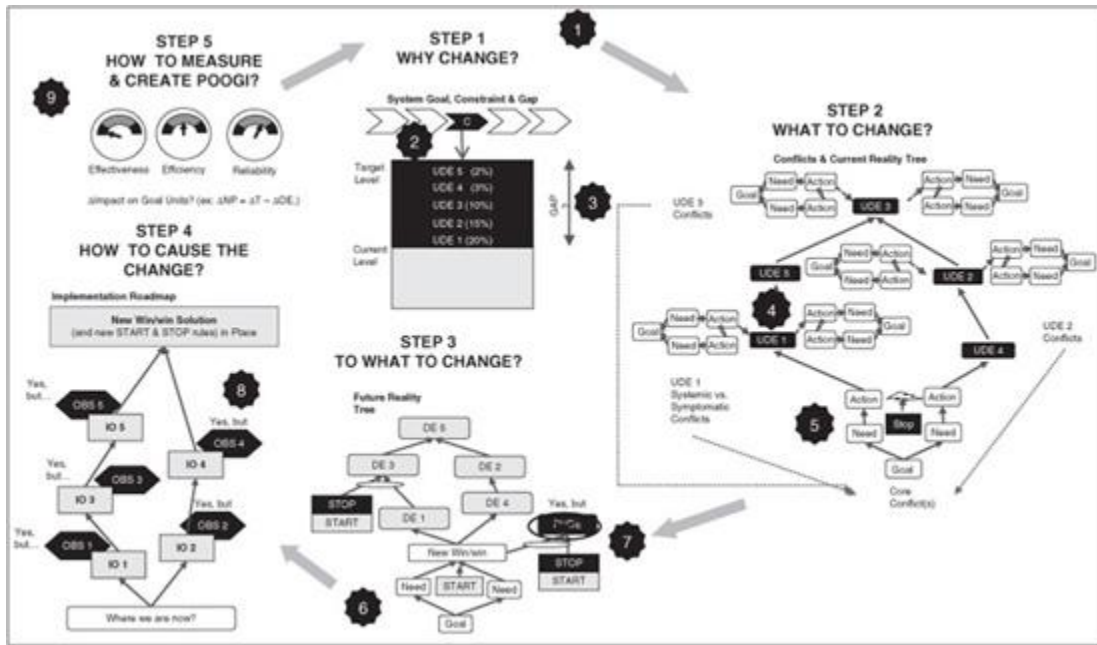


FIGURE 15-9 Barnard's new simplified TOC analysis roadmap.

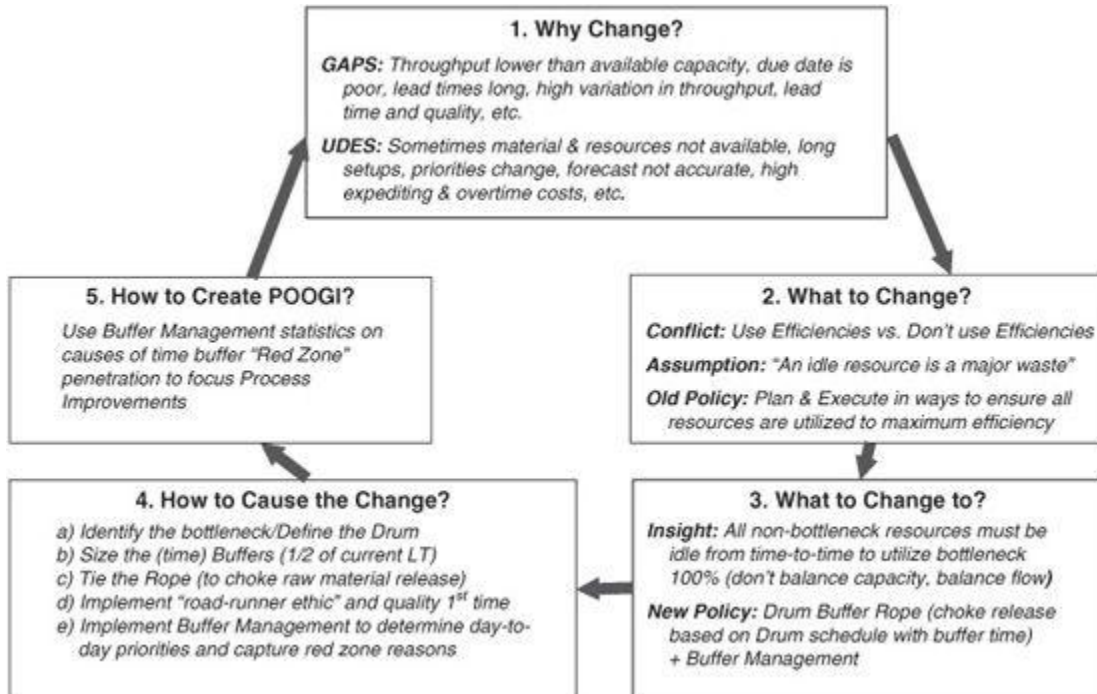


FIGURE 15-10 Applying the five questions: Managing operations the TOC way.

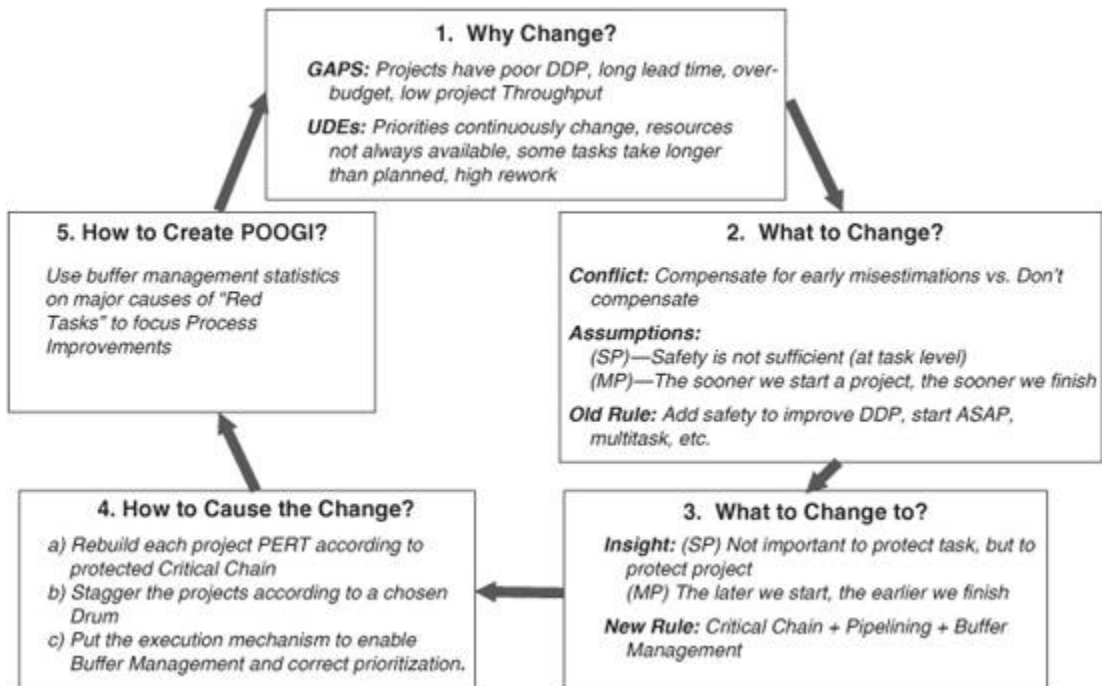


FIGURE 15-11 Applying the five questions: Managing projects the TOC way.

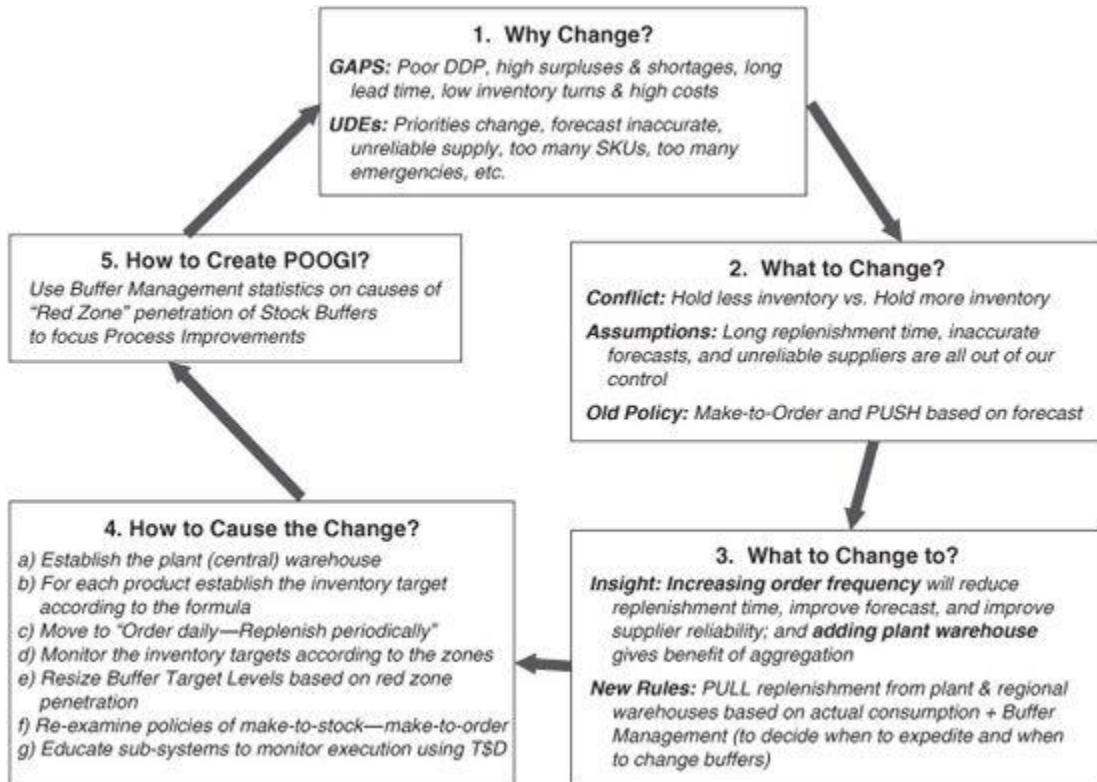


FIGURE 15-12 Applying the five questions: The TOC analysis on managing distribution/supply chains the TOC way.

3. Local efficiencies must be abolished.

4. A focusing process to balance flow must be in place.

The fourth concept, “a focusing process to balance flow (between demand and supply)” is needed to identify where to focus process improvement. Goldratt says that Ford used direct observation (of where flows were delayed), while Ohno used the gradual reduction of Kanbans (the number of containers and then the gradual reduction of parts per container) as per his famous “river flow over the rocks analogy”; reduce the water level—the Kanbans—and the rocks that stick out are the parts that have to be improved next to improve and balance the flow of products through the supply chain. For the more generic case where *space* and *stock buffers* cannot be used (i.e., where *time buffers*⁷

are needed to control flow release), the mechanisms used by Ford and Ohno have to be expanded.

Goldratt proposed two simple mechanisms to identify where further process improvements are needed to improve the flow (reduce flow time and increase flow rate) and to balance flow of products or services with demand (if the demand changed). The first “rough mechanism” (simple and fast) is similar to that proposed by Ford—simple observation of where WIP is building up in the system. The “bottleneck” that is limiting further improvements in flow is behind the WIP buildup and that is where process improvement (to better exploit scarce capacity) or capacity elevation needs to be focused to improve flow.

The second (more sophisticated mechanism) capitalizes on the buffering mechanism used by TOC. If a specific order or project task’s time buffer goes into the red, or a specific material or product’s stock buffer goes into the red or black (stockouts), then not only should this order or task be prioritized and expedited to ensure due date or availability commitments are not jeopardized, but we should also record “What (resource) the order or task was waiting for?” (Knight, 2003). On a frequent basis (e.g. weekly), these reason codes are then analyzed using Pareto to determine which resource caused most of the “reds” and “blacks.” This resource is where process improvement or elevation should be focused for the next period. This second focusing mechanism is now the recommended focusing method by Goldratt⁸ for the TOC way of managing operations, managing distribution, managing projects, and managing the sales funnel. An example of the details of the focusing mechanisms proposed by Goldratt for driving systemic CI can be found in the later section “Using the S&T to Monitor Execution.”

Lessons from CI Methods Developed by Ford and Ohno and Other Giants

Ford and Ohno were probably the first to find a systematic way to break the CI conflicts. Both Ford and Ohno ensured that a culture of continuous experimentation was put in place within each department and within every level of the organization to identify improvement opportunities that would help improve flow and reduce waste and then to encourage development and testing solutions to do processes better, faster, simpler, and with less waste. In both organizations, management was responsible for ensuring that departments focused their limited resources on those improvement opportunities that would help them realize their vision for reducing the total time from raw material to finished product with the least waste (Ford) or to reduce the total time from order to receipt of cash with the least waste (Ohno's vision for Toyota).

In Ford's (1926) *Today and Tomorrow*, he gives an indication of his CI approach when he says, "We do not make changes for the sake of making them, but we never fail to make a change once it is demonstrated that the new way is better than the old way" (1926, 53) and "our method is essentially the Edison method of trial and error" (1926, 64).

In Ohno's (1988) *Toyota Production System: Beyond Large-Scale Production* he said, "All we are doing is looking at the timeline, from the moment the customer gives us an order to the point when we collect the cash. And we are reducing the timeline by reducing the non-value-adding wastes" (1988, ix).

The other important part of the solution at Ford and Toyota, was to achieve and sustain continuous improvement was standardization of work (but ensuring the new standard would always be challenged). Ohno is famously quoted as saying (Shimokawa et al., 2009, 9), "Where there is no standard, there can be no kaizen." Without standard work,

we cannot be sure what impact our changes will have on our process and company performance.

It is clear that Henry Ford and Taiichi Ohno both approached the problem of achieving continuous and (when required or possible) step-change improvement in the same way. They started with the belief that anything can be improved, communicated a clear vision of where CI will be most valuable to the organization, and created an environment to encourage continuous experiments to find better, simpler, faster ways of doing things with less waste. They then made sure that there are continuous audits to ensure alignment (no inherent conflicts) between organizational policies. This is in full alignment with the direction of the solution proposed by TOC today.

Importance (and Risks) of Measurements and Incentives

Measurements play three important roles in CI and auditing:

1. To help managers determine the status of the system (good/bad).
2. To help managers determine the likely cause of the system status.
3. To drive the right behavior (doing what should be done) and discourage or prevent the wrong behavior (doing what should not be done) for all the stakeholders.

TOC's *Buffer Management* (BM) satisfies all three conditions as it provides a reliable mechanism to indicate the status of the system (the percent red and black within TOC's time or stock buffer status indicate to what level the system is in control or not.). The level and causes of these buffer penetrations can be used to track the level and causes of downtime or unavailability on *capacity*

constrained resources (CCRs) and level and causes of delays on the critical chain (longest chain of dependent events) to provide an indication of the likely causes of the system status.

With respect to the third role of measurements, Goldratt realized early on the important part that measurements played in the behavior of people, which drives their contribution toward organizational improvement, inertia, or decay. Goldratt's insight (Goldratt, 1990b, 145) was captured in his now famous quote "Show me how you measure me and I'll show you how I behave!" In BM, a "black" or "red" status serves as a visible signal that everyone needs to prioritize and, where possible, expedite such orders (to drive the desired behavior).

One aspect not frequently reported on is Goldratt's insight that it appears to be more important to remove "bad" measurements that drive "bad" behaviors (such as local efficiency measurements that result in local optima and poor synchronization), than it is to replace these with "good" measurements. He has also frequently warned against incentive schemes intended to motivate and drive improvements.

Why? Surely, it makes sense that when we *stop* using one measurement we should *start* using another or else we face the risk of people falling back in line with the old measurements. Surely, it makes sense that if you want people to continuously improve, you should link performance against these measurements with appropriate incentives (IF "good behavior" THEN the "Carrot" and IF "bad behavior" THEN the "Stick" consequences).

Like many of the "counterintuitive" insights of TOC, the cause-effect relationship between incentives, motivation, focus, collaboration, and how this affects the level of performance of people is quite misunderstood within most organizations. In fact, there is a major mismatch between what the social sciences have known about the effect of

incentives on performance and problem solving and how most of the incentive schemes used by organizations, work today (Pink, 2007).

Scientific research over the past 40 years has proven the “common wisdom” that incentives drive higher performance is, for a large group of boundary conditions, simply not true. Incentives, in many cases will contribute to a vicious cycle of decaying performance (or at least stagnation) rather than a vicious cycle of continuously improving performance. The first scientific research into the relationship between incentives and performance was by Sam Gluxberg, who used the “Candle Problem” designed by Karl Duncker (1903–1940) in 1926 as a way to measure how cognitive problem solving is influenced by incentives. People are challenged with figuring out how to attach a candle to a wall in a way that would prevent wax from dripping on the table ([Fig. 15-13](#)).

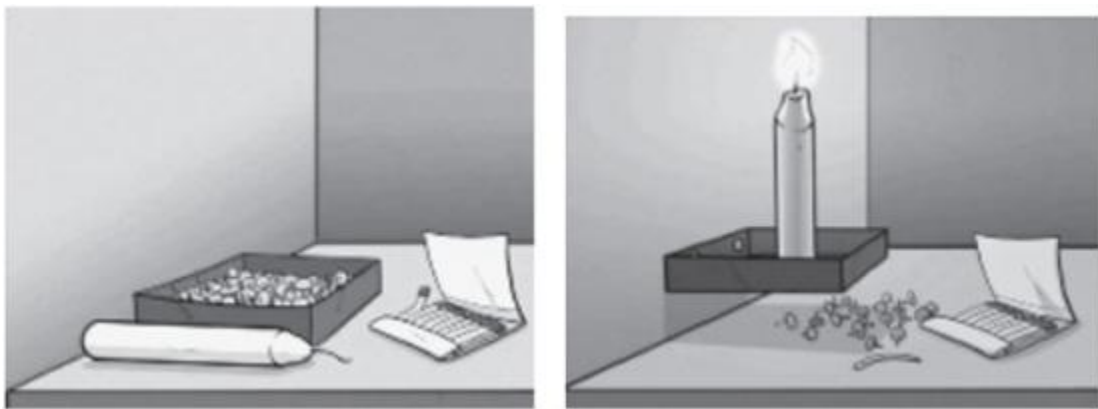


FIGURE 15-13 Karl Duncker’s candle problem to measure cognitive problem-solving skills.

Duncker found (Pink, 2007) that most people struggled due to what he called “functional fixedness”—a mental block against using an object in a new way that is required to solve a problem.” Most people eventually figure it out (to attach the box used for holding the thumbtacks with the

thumbnails against the wall to provide a base for the candle), but it takes them a while to get it. Years later, Sam Gluxberg decided to see how a monetary incentive would affect people's performance on the candle problem. He told one group if they were among the fastest 25 percent, they would get \$5. If they were the fastest in the entire group, they would receive \$20. Naturally, the people offered the incentives completed it faster, right? Wrong! In fact, they took an average of 3 minutes *longer* than those who were asked simply to perform the task as fast as possible, explaining that their results would be compared with the test standard.

Sam Gluxberg then repeated the same experiment but changed it to make the solution more obvious by placing the thumbnails next to the box, rather than inside the box. In this case, incentives fulfilled their purpose. What are the lessons from these two simple experiments?

Financial incentives tend to focus the mind and as such only tend to be productive on left-brain tasks, that is, relatively simple problems with a clear set of rules and a single solution. In contrast, when financial incentives are offered to people to solve more right-brain tasks—those problems that are more conceptual or complex in nature and require greater use of cognitive power—the incentives actually make the problem harder to solve because they narrow the focus when the solution tends to be on the periphery and so the solver needs to be thinking more holistically and laterally (thinking out-of-the-box).

These results were confirmed by an extensive study lead by Dr. Bernd Irlenusch at the London School of Business whose team studied 51 "Pay for Performance" plans inside companies and found that financial incentives can result in a negative impact on financial performance (e.g. financial incentives for sales people involved in complex sales will lower, rather than increase their success rate).

So, science has known about these flawed links between problem-solving and financial incentives for decades, and yet despite that, they endure. At the same time, more and more of the work we do is shifting to right-brain thinking as we delegate the routine, rule-based stuff to computers and outsourcing agents. But what is the solution?

Pink (2007) suggests that we move to incentives that are based on intrinsic motivators such as *autonomy* (e.g., opportunities to be independent such as Google's 20 percent "do what you want" time rule), *mastery* (e.g., opportunities to improve and excel such as Toyota's Kaizen events), and *purpose* (e.g., opportunities to be driven by what really matters to them and others in their organization). An example frequently used to prove the power of intrinsic motivators at the organizational level is how Encarta, with its teams of thousands of highly paid contributors and the backing of Microsoft, was beaten by Wikipedia, which depended on volunteers driven by a common purpose, an autonomy to contribute when and how they wished within certain guidelines and with the opportunity for mastery.

As one might expect, there are other problems with measurements and incentives. For example, when there are (many) conflicting measurements—something that frequently happens in environments that implement a balanced scorecard without aligning each measurement to a business strategy (strategy map)—people will tend to focus on those measurements they believe are most important in the eyes of management, neglecting the others (which might be more important), and making performance unpredictable. For example, if a production manager is responsible to achieve both high due date performance and their monthly cost recoveries, which they believe is their prime measurement, then it is likely that the manager will compromise on due date performance toward the end of the month to meet the targeted tons per hour for the month.

Ensuring the New Direction Addresses All Major UDEs

Overcoming the Problem of Low Expectations for Change

Previously, we identified one of the consequences of the vicious cycle in CI as stakeholders (especially top executives) having low expectations for the impact of change initiatives. To address this problem and ensure that all stakeholders have the same (high) expectations for the outcomes of the selection and implementation of *any* changes to better exploit their system constraint or to elevate it, Goldratt (2008b) recommends the adoption of the six success criteria listed in [Table 15-4](#) together with the logic of *why* each is needed, and a recommendation based on extensive field-testing, on *how* these can be used. Such extensive field-testing (Barnard, 2009) has also shown that these criteria help prevent mistakes of omission and commission in the selection and implementation of changes and that these criteria should be shared with managers and employees at all levels especially during the analysis and “buyin” phases of change initiatives and for use during ongoing audits of these initiatives.

Overcoming “Not Seeing” the Inherent Improvement Potential

The famous quotation, “Necessity is the mother of invention,” can be traced to Plato’s Republic, book II, 369C, which was written in 360 BC. We all know that crises allow us to challenge and overcome prevailing assumptions and identify and unlock potential we never knew existed. However, what if you do not have a real crisis now? In such situations, the literature on managing change is quite consistent—*good leaders should create a “crisis” by creating a large gap between the current level of*

performance and the goal. An example of this is a new CEO coming into an organization that is already doing well at 10 percent profit to sales and then (to inspire them to higher performance) gives the team the goal of doubling their profit to sales (to 20 percent) within three years.

In the case where there is no crisis yet, but where we can observe a stable or growing gap between the actual performance of an organization and its goal, we should see this as a warning and opportunity that a breakthrough is needed. We should start by answering what could cause such a gap. There are at least two hypotheses for a cause.

Hypothesis #1: The system's starting conditions (its capacity, capability, etc.) are simply insufficient to meet the demand and the only solution is to "elevate" the system constraint(s) (constraining starting condition) by investing in more resources or better resources. This hypothesis of the underlying cause for a gap is quite a common claim. "If you want my department to do more . . . I need more resources, better systems, etc."

Hypothesis #2: The system's starting conditions (its capacity, capability, etc.) are sufficient to achieve significantly higher levels of Throughput within significantly shorter lead times than currently but capacity, time, and costs are wasted due to the current mode of operations. The solution in this case will be to "better exploit" (not waste) the potential of the system constraint (i.e. always try better exploiting before elevating the system constraint).

Success Criteria	Why?	How to?
Every change must . . . Deliver EXCELLENT RESULTS	There is significant variation (noise) within the "goal units" of any system. If the variation is 10 percent and we target 5 percent improvement, we cannot measure it. In addition, there are many ways to achieve 5 percent improvement, but very few that can give 50 or 100 percent (these are the ones we target).	At the beginning of the change initiative, a very ambitious target is set for improvement in goal units. We validate that achieving the target will be measurable (outside current noise) and what the consequences on all stakeholders will be if the target is achieved or not achieved by quantifying the likely (range of) impact on ΔT , ΔI , and ΔOE .
Every change must . . . Be based on WIN-WIN-WIN for all stakeholders.	In any system with multiple stakeholders, a "lose" for one quickly degenerates to a "lose" for everyone. Where current systems are in a perceived win-lose, getting agreement that we will only accept solutions based on win-win-win goes a long way to rebuilding trust and respect.	The criteria of win-win-win are shared with all participants upfront and we get them to share stories of what happens when new solutions are perceived as a "lose" for one or more stakeholders. We end with a commitment from all stakeholders that "from today, only win-win-win solutions will be acceptable."
Every Change must . . . Be LOW RISK compared to the likely BENEFITS	Most managers, especially those in the public sector are very good at determining the risk of doing something, but not at determining the risk of not doing something and also not good at differentiating between taking calculated risks with fast feedback versus taking uncalculated risks with slow or no feedback.	At the beginning of the change initiative, the facilitator covers a few examples to enable participants to differentiate between an action that has high probability but low impact of failure and high impact of success (playing the lottery) from those actions that have low probability but high impact of failure and relatively low impact of success (e.g., playing Russian roulette) and then let stakeholders apply these criteria to the change being considered.
Every change must . . . Be SIMPLER than before	Changes that are more complex are likely to be resisted (as they are likely to require more effort), be misunderstood, be difficult to implement, unlikely to deliver quick results, and more likely to result in unplanned LDEs.	Use the quote by Albert Einstein to set expectations of simplicity "Any fool can make things . . . more complex . . . it takes a touch of genius—and a lot of courage to move in the opposite direction . . ." and let participants explain why SIMPLER is the key (easier to understand, implement, etc.) and then validate whether a proposed change is "simpler" or "more complex" than before.
Every change must . . . Be defined as ACTIONABLE information with FAST FEEDBACK	Unless changes have been defined as "actionable," we should not expect that they would be implemented. In addition, the shorter the feedback loops, the quicker we will know whether our solution is necessary and sufficient to achieve the expected results. The longer the feedback loop, the longer it will take to identify errors of omission or commission.	All change initiative stakeholders and participants are encouraged to ensure that all changes proposed will be actionable and measurable with fast feedback loops to confirm whether agreed changes are being implemented and whether expected results are achieved. To validate whether a change was communicated as "actionable information," we recommend that stakeholders explain "how I will apply/implement the changes and how I will know if it is working."
Every change must . . . Be checked so that if it really works, it will not SELF-DSTRUCT	Almost every change has potential negative consequences on one or more stakeholder if it is really successful. Not being prepared for such negative consequences can cause a successful change to self-destruct. A good example is the destruction of a reliability-based competitive edge when orders increase too fast, causing lead times to increase exponentially and reliability to decay.	All change initiative stakeholders and participants are encouraged to identify possible negative consequences if the planned change works "too well." Those that raised such reservations are asked to explain the chain of cause-effect and suggest ways to identify when such negative consequences can be triggered and what can be done to prevent them (to reduce the time to detect and correct).

TABLE 15-4 Success Criteria Recommended by Dr. Eli Goldratt (2008c)

How can we validate whether Hypothesis #1 (no significant inherent potential) or Hypothesis #2 (significant inherent potential) is most valid for a specific organization?

Let's start with the general facts (governing principles) about any system and see what we can deduce from these.

Fact 1: The system constraint (bottleneck) governs the Throughput (flow rate) of goal units for the whole system.

Implication: The system (on average) can never produce more goal units than what the constraint is capable of. However, if constraint capacity is wasted through starvation, blockage, breakdowns, or rework, then the system will achieve a lower Throughput than what the system (based on its constraint) is capable of. The level of constraint capacity wasted on starvation, blockage, breakdown, rework, etc., can be used as a reliable way to estimate whether inherent potential exists (i.e., the opportunity to do more without investing in more

resources). The capacity lost is normally between 25 and 50 percent of the available capacity.

Fact 2: The critical chain (the longest path of dependent events considering both process and resource dependency) governs the lead time (flow time) of all goal units through the system.

Implication: The parts going through the system can never go faster (on average) than the time to cover the critical chain. However, this *flow time* will be longer than the sum of processing and movement times on the critical chain when goal units traveling through the system have to wait for a resource or a decision. The level of time wasted on the critical chain due to resource or information unavailability (delays) can be used as a reliable way to estimate whether inherent potential exists (i.e., the opportunity to do the same or more within less time without investing in more resources). The time lost is normally 25 to 50 percent of the critical chain time.

Fact 3: Every system's performance (Throughput of goal units, lead time, costs, and investments) varies over time. Sometimes there is a significant variation between the best, the average, and the worst.

Implication: The "best ever" performance shows what is possible with the current starting conditions. Normally the "best ever" is achieved under ideal or crisis circumstances. The "ideal" circumstances should be turned into standard best-practice. It is in crisis situations that we become very open to "do whatever it takes," including changing the current rules (normal mode of operation) and ignoring efficiency measurements. For example, if there is a scarcity in the market, we naturally move to a "wait for the *pull*" rather than "*push* as much as you have." Why not use *pull* all the time? "Necessity is the mother of invention," but frequently these "inventions" that got us out of the crisis don't "stick" since we go back to the "way we've always done it before."

Therefore, if when we observe a significant gap and variation between the actual performance of a system and its goal, we simply need to identify:

1. How much constraint capacity (that governs the overall system throughput) and critical chain time (longest path of dependent events that govern total lead time) is being wasted (poor constraint or critical chain exploitation).
2. How much unnecessary costs or investment incurred to validate (or invalidate) the level of inherent potential (e.g., profitability) can be unlocked without any significant investment in more or better resources.

We can represent this opportunity within the model with [Fig. 15-14](#).

We can apply the same logic to validate or invalidate whether it is possible to achieve the same Throughput with less resources (truly variable costs, Operating Expenses or Investments). We can determine this through observations, studying “best-of-breed organizations,” or simply identifying all possibilities where truly variable costs, Operating Expenses and Investments are incurred unnecessarily (events such as overtime cost, emergency shipments, or unnecessarily investing in more capacity than needed because of starvation or blockage caused elsewhere in the system). Once these categories of avoidable or unnecessary truly variable costs, Operating Expenses and Investments have been identified, we can then validate whether they exist within the organization we are analyzing and, if so, to what extent they exist as a reliable way to quantify the “inherent” improvement potential. Then, tests can validate how much of this potential we can unlock without significant investments.

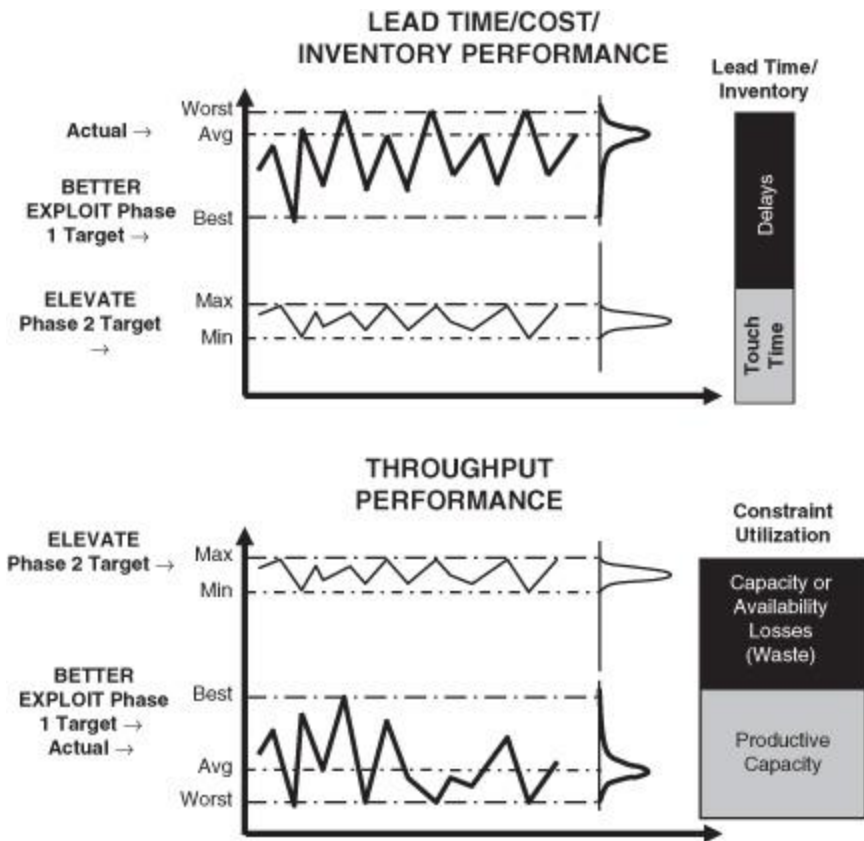


FIGURE 15-14 Quantifying inherent potential by looking for performance gaps/variation.

[Figure 15-15](#) shows a summary of the hypotheses, magnitude of inherent potential and validation that, in most organizations, it is possible to do more with less in less time. “More” by achieving higher Throughput by not wasting any constraint capacity; “with less” by achieving lower truly variable costs, Operating Expenses or Investments by eliminating the causes of avoidable costs and investments; and “in less time” by achieving shorter lead times by eliminating causes of delays on the critical chain.

Overcoming the Difficulty to Quantify the Impact of Change Initiatives

One of the key requirements of adopting a systems approach to continuous improvement and auditing is the

ability to judge the impact of decisions on the system as a whole—especially the impact of financial decisions. For most managers in organizations, the idea of trying to evaluate the impact of their local decisions or proposed investments on the “system as a whole” is a daunting, lengthy, and frequently frustrating experience (especially if they need to make a decision quickly). *Throughput Accounting* (TA) was invented by Goldratt (1990a) to meet this challenge as an alternative to cost accounting. TA (according to the IMA Statement 4HH on TOC) differs from traditional cost accounting, first in its recognition of the impact of constraints on the financial performance of an organization (i.e., if a decision impacts the constraint, the system’s Throughput will be impacted and vice versa); and second in that it separates *totally variable cost* from Operating Expenses (OE) (*all costs that are not totally variable with increased/decreased production*) to assist with faster and better decisions. This definition removes the need to allocate all costs to products and services, which frequently results in sub optimum decisions when managers erroneously assume that once OEs are allocated they become variable.

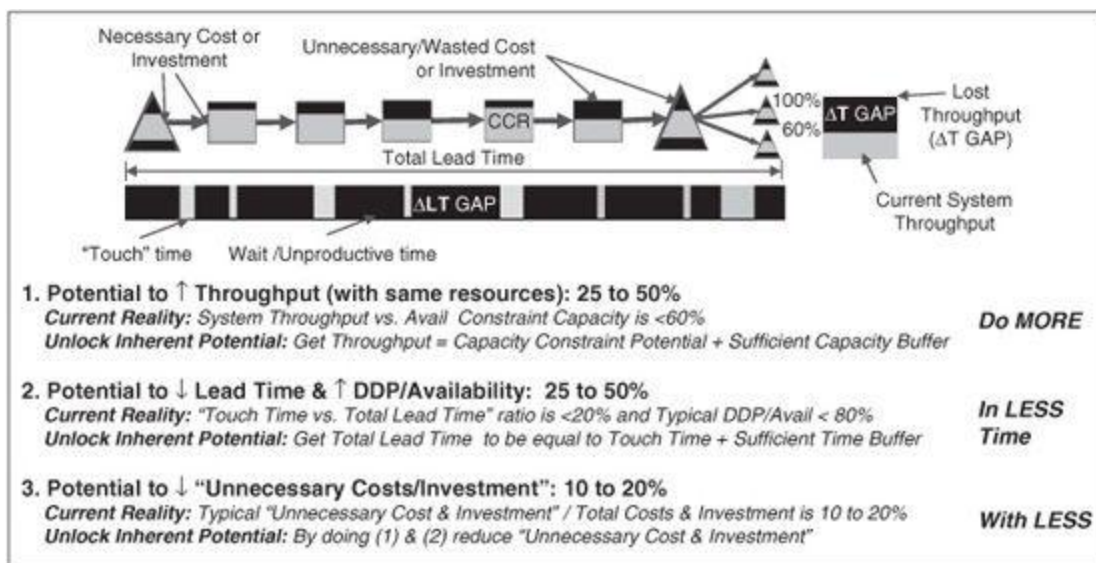


FIGURE 15-15 Identifying the inherent potential to “do more with less in less time.”

TA improves profit performance (even for not-for-profit organizations) with better and faster management decisions (Corbett, 1998), by using measurements that more closely reflect the effect of decisions on three critical monetary variables—Throughput, Investment/Inventory, and Operating Expenses (defined below). Goldratt’s alternative begins with the idea that each organization has a goal and that better decisions increase the amount of goal units that the organization can generate now and in the future. The goal for a profit-maximizing firm is easily stated—to increase profit now and in the future. TA applies to not-for-profit organizations too, but they have to develop a goal that makes sense in their individual cases. Organizations that wish to increase the attainment of the goal should therefore require managers to test proposed decisions against three questions. Will the proposed change:

1. Increase or reduce Throughput (Sales - TVC)? If yes, by how much?
2. Reduce or increase Investment (Inventory)? If yes, by how much?
3. Reduce or increase OEs? If yes, by how much?

The answers to these questions determine the effect of proposed changes on systemwide measurements:

1. Throughput (T) = Sales Revenue - Totally Variable Cost = SR - TVC
2. Net profit (NP) = Throughput - Operating Expense = T - OE

3. Return on Investment (ROI) = Net Profit/Investment = NP/I

4. TA Productivity = Throughput/Operating Expense = T/OE

5. Investment Turns (IT) = Throughput/Investment = T/I

In summary, TA is an important development in modern accounting that allows managers within both private and public sector organizations to understand the contribution of constraint resources and the frequently nonlinear impact of local actions or decisions on the overall profitability and viability of an organization.

Description	Baseline		Price Increase Impact		Volume Increase Impact		Wages Reduction Impact	
	Value	% of SR	Value	% of SR	Value	% of SR	Value	% of SR
Sales Revenue	\$100.00	100%	\$101.00	100%	\$101.00	100%	\$100.00	100%
Variable Cost	\$50.00	50%	\$50.00	50%	\$50.50	50%	\$50.00	50%
Throughput	\$50.00	50%	\$51.00	50%	\$50.50	50%	\$50.00	50%
Operating Expenses	\$40.00	40%	\$40.00	40%	\$40.00	40%	\$39.80	40%
Net Profit	\$10.00	10%	\$11.00	11%	\$10.50	10%	\$10.20	10%
Impact of 1% Change			10%		5%		2%	

TABLE 15-5 Using TA to Show the Leverage with a 1-Percent Change in Price, Volume Sold, and Wages

Knowing the impact of changes on these variables plays a vital part in both knowing *where to focus* scarce resources (especially management time) and knowing *how to predict* the impact of changes on the organization's profitability/viability. As an example, [Table 15-5](#) provides a baseline case, that shows the leverage achieved in a 1-percent increase in average selling price, a 1-percent

increase in volume sold, and a 1-percent reduction in wages on the Net Profit of the organization (10%, 5% and 2%).

Preventing and Correcting Errors of Omission and Commission

It was stated previously that two of the most common types of management mistakes include errors of omission and errors of commission. Another way to look at these errors is by relating the errors to whether action was taken based on a tested or untested hypothesis.

1. Errors of Commission: Doing what should not be done or acting on an untested hypothesis.
 - a. Do work that is not important (what Goldratt calls choopchiks⁹) or urgent.
 - b. Do the wrong thing (it will not solve the problem and may even make things worse).
 - c. Do too many things at the same time.

2. Errors of Omission: Not doing what should be done or *not acting on a tested hypothesis*
 - a. Don't act because "we still have time" or "we don't have the time (to act now)."
 - b. Don't act because "we are different (it will not work here)."
 - c. Don't act because "they will never agree (and without their agreement it's a waste of time to even start)."

Using the classification of "acting on an untested hypothesis" and "not acting on a tested hypothesis," is also useful as it helps to identify the simple solution to prevent both mistakes. To prevent errors of omission and commission we should *always check (our hypothesis) before we act/don't act, unless acting is the only way of checking.*

But, how do you check your hypothesis? It starts by recognizing that every decision we make and every conclusion we reach and communicate or act on that

contains the word “because” or “will result in” contains a hypothesis, such as “We should not make the offer because customers will never agree,” “Customers are buying less because our competitors have reduced their prices,” etc.

We have two ways to check a hypothesis. One is through logic, using the effect-cause-effect method to identify and then validate predicted effects. The more predicted effects are validated, the more valid your hypothesis, considering that even one predicted effect that is validated (if it could not happen by “fluke”) could be enough to validate. At the same time, even just one predicted effect that cannot be validated could invalidate your hypothesis (e.g., the observation of just one black swan can invalidate the hypothesis based on thousands of observations that all swans are white).

The second method, frequently referred to as the scientific method of “trial and error,” is to test the hypothesis by acting. Leonardo da Vinci said, “I am always impressed with the urgency of doing. Knowing is not enough; we must apply. Being willing is not enough; we must do.” It is only once we apply that we can really test the necessity and especially the sufficiency of our hypothesis (new solution). That is why our “injection” to reduce errors of omission and commission includes “. . . unless acting is the only way of checking.” This realization is critical to overcoming disagreements or fear. For example, it might be that the only way of really checking how customers will react to a new offer or product is by actually presenting this offer. Of course, such “tests” should be designed as an experiment and reviewed with the same rigor rather than wasting time arguing in the boardroom whether they will like it or not, or whether to first do more market research (what Goldratt calls “just a sophisticated way of procrastinating”).

Overcoming the Fear of Uncertainty

In studying the success stories such as Toyota, Walmart, and GE, it is noticeable that each of these had a leader/leadership team willing to take the responsibility to decide which philosophy or methodology to use rather than leaving it up to levels 3, 4, or 5 to decide. Not only did they decide on the philosophy, but also on the vision and made sure the connection between the two was clear to every level and every function in the organization and then empowered everyone to contribute (within the boundaries of the philosophy) to achieving the vision or goal. These leaders also showed that they are continuously challenging their own patterns of thinking and want those around them to do the same—not randomly, but in a systematic process using the scientific method.

Why is this important? It helps overcome the fears and prevent the mistakes related to the decisions and conflicts of when to change, what to change, and how to cause, sustain, and continuously improve on the change.

What can organizations do who find themselves almost paralyzed by the fear of changing (due to a culture that seems to punish failure and not recognize the courage to invent and test new ways of doing things) unless competitors made it necessary for it to do so?

This deficiency in organizations can be eliminated by taking the following steps (Barnard, 2001; Ackoff, 2006).

1. Record every important decision, including the ones *not to do something* because of reasons such as “we still have time” (when you’ve run out), “the cost or risk of doing is too high” (ignoring the cost or risk of not doing), etc.
2. The Decision Record should include (a) the event that triggered the need for change, (b) the expected effects of the decision and by when they are

expected, (c) the assumptions on which the expectations are based, (d) the inputs to the decision (information, knowledge, and understanding), and (e) why the specific decision was made (the logic) and by whom.

3. Monitor the decisions to detect any deviation of fact from expectations and assumptions. When a deviation is found, determine its cause and take corrective action (to reduce the time to detect and correct mistakes).
4. The choice of a corrective action is itself a decision and should be treated in the same way as the original decision; a Decision Record should be prepared for it. In this way, one can learn how to correct mistakes; that is, learn how to learn more rapidly and effectively. Learning how to learn is probably the most important thing an organization or individual can do.

The decision by an organization not to adopt systems thinking/holistic approach of TOC should be treated in this way. Making explicit the assumptions on which such a decision is based and monitoring them can lead to a reversal of the decision in time.

Potential Negative Branches and How to Prevent Them

Fear That "Focusing" Will Jeopardize Continuous Improvement Culture

There is a fear that focusing all scarce resources on a few high-leverage improvement opportunities will jeopardize the establishment of a culture of CI and even result in inertia or complacency for those not directly involved with the few high-leverage changes.

Organizations such as Toyota and Walmart have shown that it is possible to encourage everyone to contribute continuously to improving the processes for which they are responsible unless the proposed change to improve a local process will adversely affect organizational performance (in which case it will not be approved) or where the proposed change will require scarce resource allocation (in which case the change initiative should be planned and then “pipelined” until the necessary execution resources are available).

We should continuously improve all areas where local improvement will enable an improvement of the organization as a whole. For example, Toyota is continuously encouraging all employees at all levels to find ways to reduce waste, as long as the reduction will help reduce the total time from receiving an order until the cash is received or will reduce unnecessary truly variable costs, Operating Expenses or Investments, without jeopardizing Throughput, total lead time, or quality.

Fear That We Will Have Too Much/Too Little Because of a Change

A key part of preventing “local optima” and of ensuring that “non-constraints” do not become bottlenecks (especially during times of growth) is the understanding of the concept of “good enough.” The normal claim is that if you want more Throughput, better performance, etc., then we need more resources. All managers should ensure that their areas perform at a level of good enough (to meet stakeholders’ commitments), which means allocating not too much but also not too little resources, time, or stock. TOC recommends that the best way to establish this “good enough” threshold is to carefully manage the stock, capacity, and time as buffers needed to protect organizational performance and to monitor the status of

these buffers. If the buffer goes into the black (e.g., no stock, no available capacity, or past due) or goes into the red too frequently, it is a sign that the “buffer” is too small and needs to be increased. At the same time, if the buffer never goes out of the “green zone” (typically one-third of the total buffer size), then it is not only safe to reduce the buffer, but necessary to prevent waste.

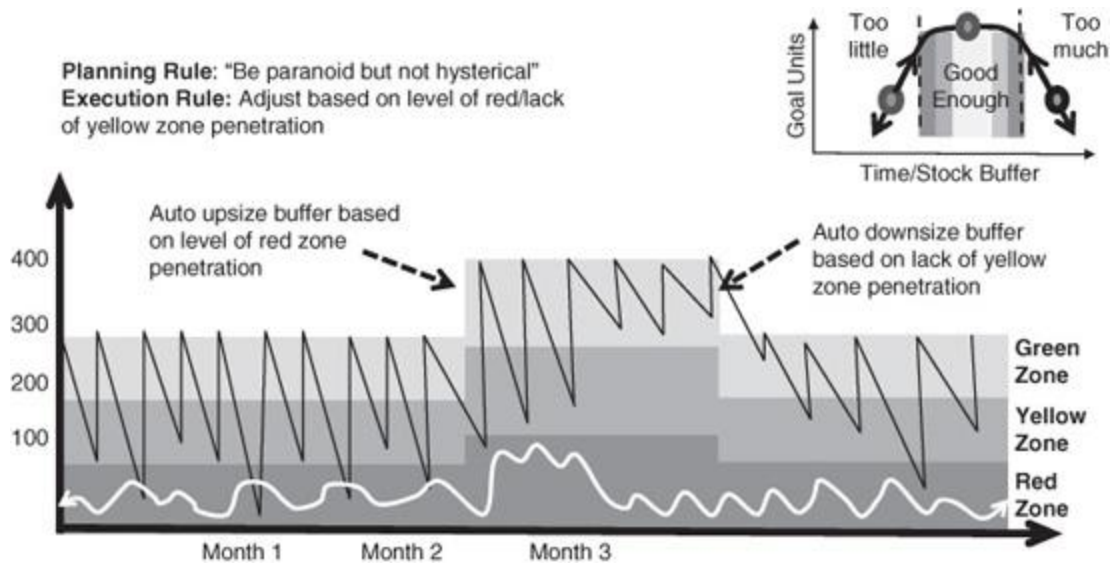


FIGURE 15-16 Using buffer penetration to maintain “good enough” levels of stock.

An example of this mechanism to maintain the right stock levels (not too much and not too little) within the TOC for distribution solution is illustrated in [Fig. 15-16](#).

This same principle can be applied to establishing the “good enough” level for any local capacity, stock, time, or even cash buffer needed to protect and improve organizational performance. Start by putting in place a starting value using the “Be paranoid but not hysterical rule” (e.g., we need three designers to meet current demand). Then establish the buffer zones, normally using one-third for each zone (e.g., 1 person = each zone), and then simply monitor the penetration of the buffer. If

frequently all three resources are utilized on important and urgent work, then it would be prudent to add more resources. If frequently two or more resources are not needed on important and urgent work, it probably is safe to reduce the team without jeopardizing Throughput. Of course, with a TOC mindset, we should always check whether excess capacity could not be turned into a competitive advantage (to secure higher prices or more sales) before considering reducing the excess capacity.

Summary of “What to Change to?”

[Figure 15-17](#) summarizes the main direction of the solution to break each of the three continuous improvement conflicts. Organizations can use TOC’s BM to identify when to change (when buffers have significant red or black zone penetration) and when not to change (buffers remain in the green or yellow). Organizations can use the 5FS of TOC together with the relevant TOC TP to differentiate between all the many parts that can be improved from the few that must be improved (now). Finally, organizations should ensure that all change initiatives are pipelined based on available capacity and that changes that do not achieve the required rate of improvement should be stopped or modified to get back on track.

Injections to Continuous Improvement Core Conflicts

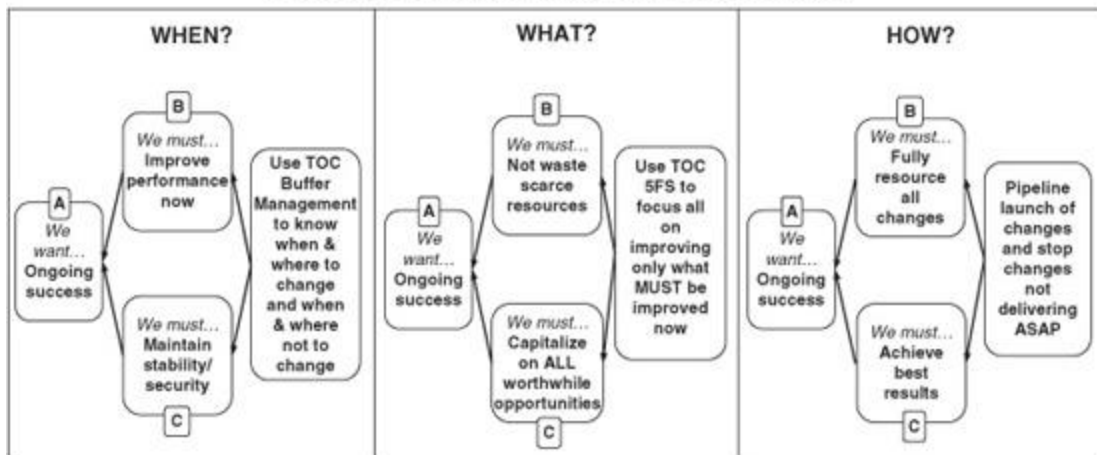


FIGURE 15-17 The TOC injections to the when, what, and how CI conflicts.

How to Cause the Change?

Typical Implementation Obstacles and How to Overcome These

Why More Organizations Are Not Adopting a Systems Approach

The logic and results that have been achieved by organizations and individuals that have used TOC's systems approach or systems thinking are compelling. However, we should ask, "If this way of thinking and this way of improving organizations is as good as they say it is, why don't more organizations use it?" Well, the typical response to this question from systems thinking/TOC experts normally includes an answer like ". . . because most organizations naturally resist change." In addition, if asked, "How do you know that this is the main reason?" they are likely to answer, "We know most organizations resist change

because most organizations have not yet adopted systems thinking/TOC.” This answer is, of course, a tautology in the same way as saying that the reason why customers use our products is because they like them and then explaining that the way we know that customers really like them is because they use them!

Russell L. Ackoff (1919–2009, an American organizational theorist and Professor Emeritus of Management Science at the University of Pennsylvania, Wharton School of Business, and one of the pioneers in the field of systems thinking) was one of the first authors in the field to try to answer the question: “Why do so few organizations adopt systems thinking?” (2006). In the article, Ackoff stated that he believes there are two reasons—one general (why organizations in general do not adopt any ideas that are considered paradigm shifts) and one specific (specifically why organizations find it difficult to adopt systems thinking). He claims that in most organizations (and in the minds of most people) failure or making mistakes are a bad thing. People would rather make errors of omission than errors of commission. This tendency is then perceived as “resistance to change,” but it really relates more to the fear of the possible failure (or even lack of success) of a proposed change. Moreover, considering how many of the past changes in organizations were considered failures or were stopped before completion, it makes sense that most people will feel it is safer to not (fully) support any new change, especially ones considered “radical.”

The specific reason, Ackoff claimed, that blocks organizations from adopting systems thinking is simply that there is a knowledge and experience gap especially on the “how to.” Most systems thinking (and probably most TOC) education programs cover only the “what” and “why” but not the detailed “how to.” Giving students and managers theories and success stories without showing which of their specific assumptions about reality will have to change and

not providing practical ways (methods, tools, systems, etc.) to apply these theories make them impractical. This does not mean that most of the required logic, methods, and tools don't already exist. It simply means that education and general awareness, yet again, lag the innovation and field experience that is being obtained by the growing number of practitioners of systems thinking (McDermott and O'Connor, 1997) and growing number of TOC practitioners.

Both the general reason and the specific reason postulated by Ackoff provide a reminder that any change initiative should ensure not only that the principles behind the change are well communicated and understood, but also that these principles are backed up by practical methods (the how to) and concrete examples of how to apply the principles on a daily basis. Without such "actionable information" and a culture that encourages experimentation and ongoing analysis of both successes and failures to continuously improve our understanding of cause-effects (what Senge [1990] called "creating a Learning Organization"), we should expect that the adoption of changes, especially radical changes or counterintuitive changes like TOC and systems thinking, will be resisted.

Enterprise Resources Planning System Does Not Support POOGI Process Such as Buffer Management

Enterprise Resources Planning (ERP) systems do not normally have functionality to support BM (recording reasons of why orders go into the red or black and providing Pareto analysis on the most significant contributors to delays or lost Throughput). However, it has been shown (e.g., Barnard, 2009, using SAP at ABB) that it is possible to relatively easily build this functionality using standard reporting tools included in most ERP systems if the company is using the TOC solution for Production (DBR/S-DBR and BM). However, if the company were using *Critical Chain*

Project Management (CCPM) to manage operations, third-party CCPM software with BM functionality would be required.

Using TOC to Focus and Accelerate Lean and Six Sigma Initiatives

The methods, mindsets, and lessons learned from the failures and successes of TPS/Lean, TQM/Six Sigma, and other improvement philosophies in achieving and sustaining CI have been in the public domain for a long time. And although there are many that successfully applied these principles and methods to their organizations, there is still a large failure rate of change initiatives, even when done under the auspices of a company-wide Lean, Six Sigma, or Balance Scorecard implementation. There are three types of mistakes frequently made in using improvement methods such as Lean or Six Sigma:

1. Selecting the largest local variation, performance gap (actual versus potential), or largest time delays (waste) rather than those with the largest impact on overall organizational performance for improvement.
2. Selecting easy, solvable problems (to achieve local efficiency, incentives, or certification).
3. Selecting problems that match our toolkit (TOC also suffers from this).

The primary departure of TOC in dealing with these three mistakes (from the continuous improvement mechanisms of TPS/Lean and TQM/Six Sigma, etc.) is that TOC does not support CI of every process. It recognizes two basic characteristics of complex systems:

1. Since constraints and non-constraints exist, more is not always better. Improving a specific non-

constraint's performance above the level of "good enough" (i.e., the level of protective capacity required to ensure it does not negatively affect the constraint) will not only waste scarce resources, but can also even result in system performance decay.

2. Due to the nonlinearity of the interdependencies, a small local delay, variation, or gap in one part can cause a large system delay or performance variation or gap and vice versa.

Therefore, TOC puts the focus on continuous improvement of the system by focusing improvement on the constraint and those non-constraints that are performing at the level below "good enough," that is, resulting in poor exploitation or lack of elevation of the system constraint. The TOC-focusing mechanism of tracking those parts whose performance is causing most of the "black" or "red" buffer statuses will ensure the excellent tools of TPS and TQM aren't wasted on improving all parts that can be improved, rather than the few that must be improved to achieve more goal units now and into the future.

No wonder that more and more of the organizations that previously used only Lean or Six Sigma or a combination, are now sharing their intentions or results already achieved in using TOC as the primary focusing mechanism for their improvement initiatives¹⁰ and as the mechanism to prevent the three mistakes listed previously.

As an example, Sanmina SCI, a large semiconductor manufacturer in the United States reported (Pirasteh, 2007) that it did a study on 21 of its factories over a 2.5-year period to measure the results achieved using only Lean (4 plants), only Six Sigma (11 plants), and TLS (Theory of Constraints as the focusing tool for Lean and Six Sigma) (6 plants). The results were staggering. Although each plant contributed toward the total savings, the six TLS implementations contributed 89 percent of the total

savings. This meant that the results of a Lean or Six Sigma implementation could be improved by a factor of 15 if focused in the right area (on the system constraint or non-constraints that limit better exploitation or elevation of the system constraint).

Other examples of the leverage that can be achieved by implementing TOC after Lean and Six Sigma include the following:

- **Delta Airlines** (Mays and Adams, 2007) reported that despite implementations of Six Sigma (1999) and Lean (2000), they filed for [Chapter 11](#) bankruptcy. In 2006, they started with their TOC implementation within their Maintenance and Repair Organization and within less than 12 months increased Throughput by 25 percent, reduced *turn-around-times* (TAT) by between 10 and 40 percent and increased revenues by 20 percent, without an increase in OE, making a significant contribution to the successful turnaround of Delta.
- **Warner Robins Air Logistics Center** reported that after a successful Lean implementation they reduced their TAT on the maintenance of the C5-Galaxy from 390 *days* to around 240 *days* between 2000 and 2005. After implementing TOC, they managed to reduce it to 170 *days* within less than 6 months (another 30 percent reduction).
- **Seagate Technologies** (Zephro, 2004) had completed over 4700 Six Sigma projects, reported \$1.2 billion in savings, and trained over 8000 employees as Green, Brown, and Black Belts by 2004. However, a review of their Six Sigma projects and feedback from stakeholders made them realize that the DMAIC¹¹ analyze stage lacked a strong effect-cause-effect tool, the DMAIC improve stage didn't have an effective tool for solution development, the DMAIC improve stage didn't have a

method to resolve conflicts, generally projects took too long to complete (average 6 + months), and Belts did not have a way to prioritize and mine projects. To solve these problems, Seagate introduced the TOC TP to supplement the DMAIC tools. The results have been impressive: Project approvals increased since they are more focused and they have more clearly stated problems and solution paths that include how all stakeholders will benefit. They also reported less resistance to buy-in for proposed solutions and faster and more successful project completion (project completion rate increased by 80 percent, number of projects completed increased by 70 percent). Graduating Transactional Belts voted *Current Reality Tree* (CRT), *Future Reality Tree* (FRT), and *Evaporating Cloud* (EC) the most useful Six Sigma tools in every wave since introduction.

Not only TPS/Lean and TQM/Six Sigma are benefitting from the focusing mechanism of TOC. Researchers, using other improvement methods such as Business Process Reengineering (Al-Mashari et al., 2001), Agile Management for Software Development (Anderson, 2003), Activity-Based Costing (Vergauwen and Kerckhoffs, 2009), ERP Software Implementations (Barnard, 2001; 2009), and Balanced Scorecard (Breyfogle, 2008), are starting to present case studies on how the use of TOC as the “systems or holistic approach” enables achievement of faster and more significant improvements and reduces or prevents the past high-failure rates of powerful methods wasted on local optima. These results show to what extent TOC as a focusing mechanism can unlock inherent potential.

Using TOC's S&T as a CI and Auditing Tool

Introduction to TOC's S&T TP

The introduction by Dr. Eli Goldratt and coworkers (2002) of a “new” TOC-based TP called a *Strategy and Tactic (S&T) tree* is being viewed by more and more executives and managers who have been exposed to it, as one of the most important breakthroughs in ensuring that holistic business or organization strategies are defined, properly validated, communicated, and implemented to achieve harmony within organizations. The S&T tree, they believe, can for the first time provide them with a practical process and logical structure for defining and communicating all the necessary and sufficient changes as well as the sequence of implementation of these changes to achieve more goal units for the organization—not just what to change but, more importantly, what not to change, and especially how to implement the changes and why.

As with many breakthroughs, this breakthrough started with a simple question by Goldratt (Goldratt et al., 2002): If “strategy” is really at the highest level of an initiative or organization and defines the direction that dictates all activities, and “tactics” are lower down in an initiative or organization and define the activities that are needed to implement the strategy, then where does “strategy” end and “tactics” begin?

Goldratt realized that answering this question required that the words “strategy” and “tactic” had to be defined more clearly than before. His new definitions were inherently simple, yet powerful. He decided to define “strategy” as, simply, the answer to the question: “What for?” (The answer is the objective of a proposed change.) “Tactic” is defined as, simply, the answer to the question “How to?” (The answer is the details of the proposed change.) From these definitions, it is clear that every

strategy (What for?) should have an associated tactic (How to?) and therefore strategy and tactic must always exist in “pairs” and must exist at every level of the organization.

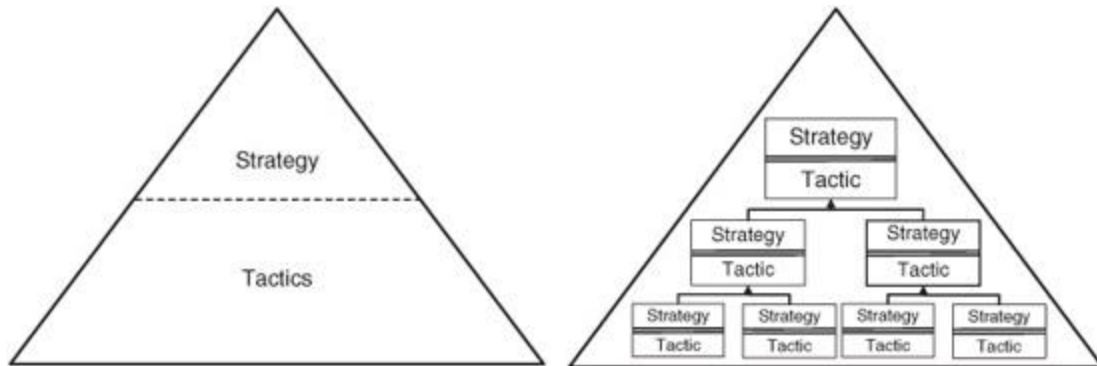


FIGURE 15-18 Traditional versus Goldratt’s definition of S&T.

[Figure 15-18](#) shows the implication of the changes to the definitions of strategy and tactics proposed by Goldratt.

Goldratt’s proposed Strategy and Tactic tree (Goldratt et al., 2002) therefore can be viewed as simply a logical tree of the proposed changes that should be both necessary and sufficient to ensure the synchronized achievement of more goal units for the organization. However, any logical tree is only as valid as the assumptions on which it is based. Therefore, it is the responsibility of managers at every level in the organization to not only contribute to defining and communicating the strategy and tactic for each proposed change, but also to define and communicate the logic of the proposed change. This includes, why the proposed change is really necessary to achieve the higher level objective and ultimately the goal of the company, why they claim it is possible to achieve the objective (strategy) of the change (especially considering it has probably never been achieved before). They must then address why they claim their proposed change (tactic) is the best or even the only way of achieving the strategy of the change, and finally, the advice or warning they would give to their subordinates to ensure

the sufficiency of the implementation of the proposed change.

Each S&T node in the S&T is therefore simply a proposed change that should answer:

1. Why the change is needed (Necessary Assumption)?
2. What the specific measurable objective of the change is (Strategy)?
3. Why we claim the strategy is possible and what specific requirements, potential negative branches, or obstacles must be considered when selecting from the alternative ways (tactic) for achieving the strategy (Parallel Assumptions linking Strategy with Tactic)?
4. How to best achieve the objective of the change (tactic); for example, what changes should be made to processes, policy, IT systems or measurement?
5. What advice or warning should be given to subordinates, which, if ignored, will likely jeopardize the sufficiency of the steps they would take to implement this tactic and which is likely to be ignored (without the warning) (Sufficiency Assumption)?

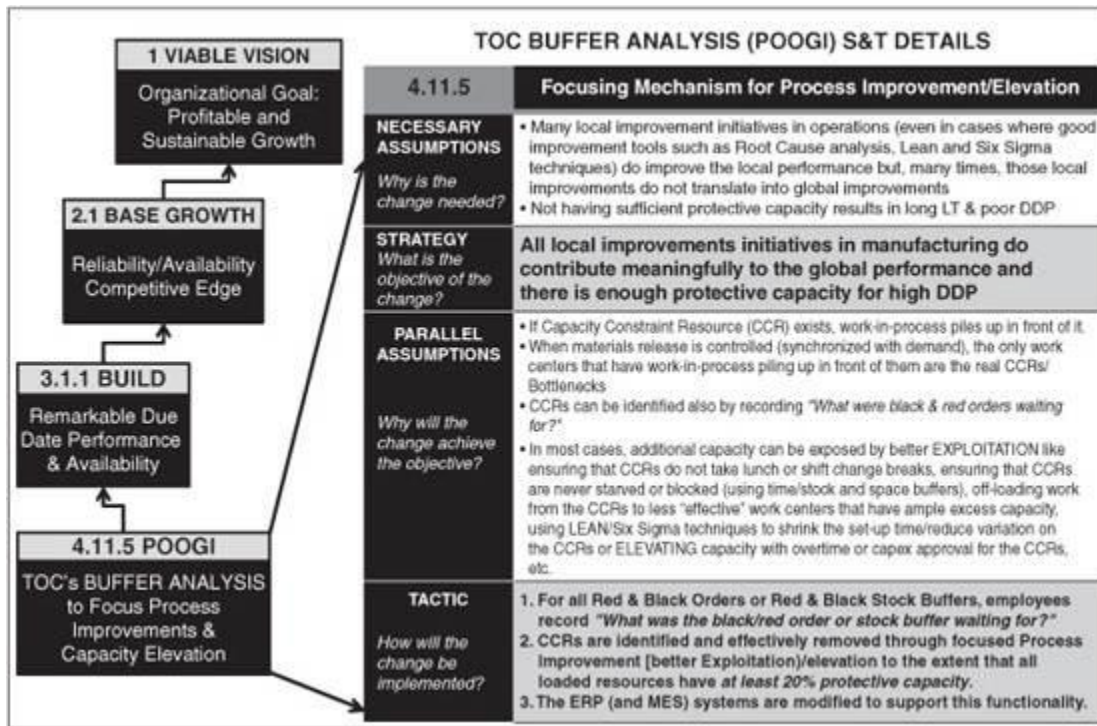


FIGURE 15-19 S&T level 4.11.5 for achieving a process of ongoing improvement.

[Figure 15-19](#) shows how the S&T can be used to define and communicate the necessary change to “Achieve a Focusing Mechanism for Process Improvement/Capacity Elevation” by explicitly answering the questions of:

1. Why the change is needed (necessary assumptions),
2. What the specific measurable objective is for the change (strategy),
3. Why the objective is possible and why the tactic is the “best” way (parallel assumptions), and
4. How (the new procedure for implementing the change).

The most recent update on the applications of TOC’s BM as a POOGI can be found in the generic S&T trees¹² released

by Goldratt. The relevant POOGI parts of the S&Ts for manufacturing companies, distribution companies, and projects companies are included in these generic S&T trees. Each of the POOGI processes typically have three components: (1) recording reasons for red and black buffer status, (2) conducting a Pareto analysis on reasons for identifying the primary causes of delays/unavailability, and (3) developing and implementing improvement projects (e.g., Kaizen events) to address the primary causes of delays/unavailability.

No.	Engines of Disharmony	Engines of Harmony
1	Not knowing my own contribution or thinking that my contribution is not recognized.	Knowing exactly how I should contribute and knowing that my contribution will be recognized.
2	Not knowing how others should contribute or thinking that their contribution is recognized too much or too little.	Knowing exactly how others will contribute and knowing that their contribution will be recognized.
3	Inherent conflicts on which rules to use to best contribute to the company goal.	All "rules" are aligned with the goal and strategy of the company.
4	Unresolved gaps between responsibility and authority (resulting in ongoing fire fighting).	Gaps between responsibility and authority are systematically identified and removed.
5	Inertia or fear that is not verbalized, which block CI.	Constraint focused CI process and culture is in place.

TABLE 15-6 Engines of Disharmony versus Engines of Harmony

Using the S&T Tree to Achieve Harmony within Organizations

How much of this knowledge is defined properly, documented, communicated, and systematically validated/invalidated within a typical organization? Field experience shows that very few organizations, if any, have taken the time to develop their business strategy and tactic to this level of detail. Unfortunately, there is a price to pay for managers and employees at all levels for not being able to answer these simple questions for their area of

responsibility. The price you pay is the risk of what Goldratt, in recent public events, has called “disharmony” in the organization. Goldratt identified five “engines of disharmony” that make achievement of CI culture and harmony within any organization difficult and to which a well-defined and communicated S&T can provide the engines of harmony ([Table 15-6](#)).

Using the S&T to Monitor Execution

Edison’s famous quote, “Vision without execution is a hallucination,” is a reminder of the simple fact that without follow-up and follow-through, we will not get the desired results.

Since the S&T contains all the objectives that have to be achieved (strategies or “know-what”), and the necessary changes that have to be implemented (tactics or “know-how”) at each level and within each part of the organization, as well as all the related assumptions (“know-why”), the S&T can be used as a primary auditing tool. Are we achieving our objectives? Have we implemented the agreed tactics? If not, which of the assumptions that were made are no longer valid and how can or should this be corrected?

Using the S&T to Identify and Systematically Remove CI Conflicts

Anyone who has tried knows the challenges in creating an S&T. There is the challenge of knowing which questions to ask to identify the necessary and sufficient changes and the sequence in which these changes must be implemented, the challenge of answering these questions (using solid cause-effect), and finally the challenge of verbalizing the answers in a way that will ensure that the proposed changes are communicated clearly as “actionable information.”

However, can we use the other TOC TP to help overcome some or even all of these challenges?

Over the past two years, a new process to use traditional TOC TP (such as UDEs and the Conflict Cloud) has been tested as part of a CI and auditing process to validate or even create new S&T blocks (Barnard, 2009). [Figure 15-20](#) shows this new process applied to managing projects. The process can be initiated at any level within the organization where a clear gap in performance exists that currently limits the achievement of the higher-level goal for the organization.

Step 1 involves identifying this performance gap, validating the extent and consequence of not closing this gap in relation to the goal of the organization (e.g., using the impact of the gap on the organization's Throughput, Investment, and Operating Expenses), and finally identifying the major UDEs (independent causes that contribute to the gap). In the example that follows, these UDEs reflect the six most common causes of projects being completed late, over-budget, or under scope. Step 2 involves defining the subordination conflict that blocks removal of each of the most significant UDEs (in terms of their contribution to the performance gap). The example shown in [Fig. 15-20](#) shows the subordination conflict related to one of these UDEs. Step 3 and Step 4 involve identifying the erroneous assumptions that cause the conflict and identifying the new assumption and related new rule that will break the conflict. Finally, Step 5 converts these insights into the structure of the S&T with the objective of the conflict (A) being equivalent to the higher-level objective of S&T (e.g., "Meeting all project promises"), the necessary conditions (B&C) being equivalent to the strategy, the assumption that was challenged being equivalent to the parallel assumptions, and the new injection (the new rule) to satisfy B and C equivalent to the tactic. The original UDE (bad multitasking)

and its consequences are equivalent to the necessary assumption of that S&T block.

Summary of How to Cause the Change

The implementation of a new TOC-based POOGI can be blocked by a number of implementation obstacles. Significant obstacles that will have to be overcome are the typical resistance to change and lack of actionable information, both of which can be addressed by using TOC's buy-in process and generic S&Ts (which describe the recommended change in detail). Another significant obstacle is the organization already investing in another CI methodology such as Lean, Six Sigma, or Balanced Scorecard. A number of organizations have already presented how they have integrated TOC as the focusing mechanism into their existing CI methodology. Another implementation obstacle is the fact that most ERP systems do not (yet) support TOC's BM (TOC experts consider BM a necessary component of all logistics applications), but again, it has been shown that standard ERP systems such as SAP can be modified to provide this functionality. The last implementation obstacle we reviewed is what Goldratt calls "the engines of disharmony." Goldratt proposed that the new TOC TP, the S&T, is the best way to remove these engines of disharmony and we can also use the S&T as a primary auditing tool.

Summary of Continuous Improvement and Auditing the TOC Way

The efforts to improve organizations continuously have resulted in the fact that change initiatives have become a continuous journey in most organizations today. Over the

past 100 years, there have been major advances in the development of the mindsets and methods needed to continuously improve and audit organizational performance. Unfortunately, despite these advances, the majority (60 to 80 percent) of change initiatives still fails to meet their original objective and some even cause decay in performance or failure. This high-failure rate frequently triggers a vicious cycle of higher resistance to change and lower expectations that result in many new initiatives not receiving the required support and resourcing that, in turn, increases the probability of failure, which again will result in higher resistance to change and low expectations. But why do people resist change? There are at least five different reasons, each requiring a different solution to overcoming the resistance to change. We resist change when:

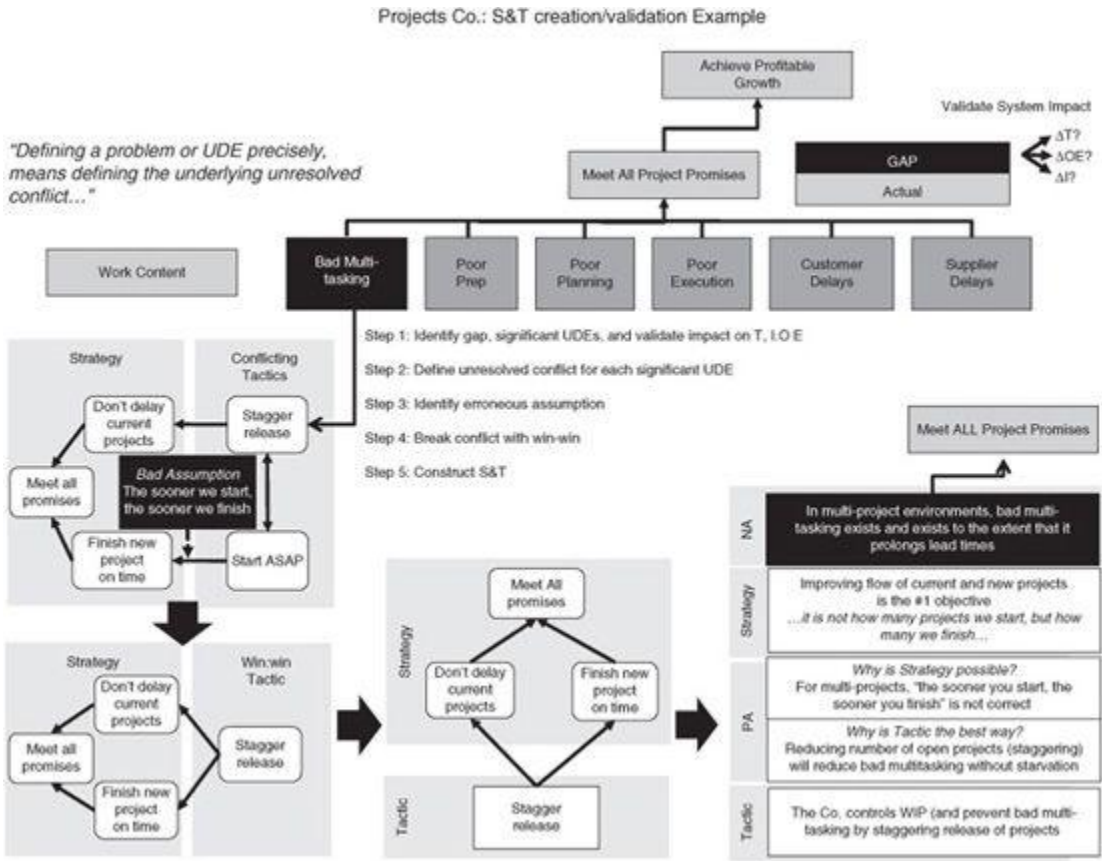


FIGURE 15-20 Proposed process using conflict cloud to creation/validate new S&T entities.

1. The change (required to solve a problem or break the vicious cycle) is counterintuitive.
2. The change has no perceived benefit or the perceived “cost” outweighs the benefit (to us).
3. The change is not detailed sufficiently to provide actionable information.
4. The change has potential negatives to us or other stakeholders.
5. The change has implementation obstacles.

The mistakes made as a result of this resistance can be classified into errors of omission (not doing what should be done), errors of commission (doing what should not be done or doing the right thing not in the right way), and errors of detection and correction (not detecting or correcting or taking too long to detect or correct despite available data). To prevent these “errors,” organizations need to adopt a holistic focusing mechanism to identify *when* to make changes (and when not) *what* changes to make (and what not) using TOC’s 5FS and TOC’s TP, and *how* to implement these changes in a sustainable way. Organizations also need a decision-support mechanism to judge the impact of decisions on the system as a whole (e.g., TOC’s TA) as well as a fast feedback mechanism to reduce the *time to detect* and *time to correct* for mistakes made (e.g., TOC’s BM). Last, organizations that want to achieve CI have to create a culture that will ensure that fear of failure does not become a major cause of errors of omission and errors of detection and correction.

Henry Ford and Taiichi Ohno provided great examples of how to create such a holistic CI mechanism and the necessary culture (based on the scientific method) to inspire CI at all levels and within all functions but which focused on supporting the overall vision. They both shared a belief that anything can be improved and made sure this belief, together with their vision of where CI would be most valuable to the organization (and other stakeholders), was continuously communicated and practiced throughout the organization. They also knew the importance of creating a safe environment (to reduce errors of omission caused by fear of failure) to encourage continuous experiments to find better, simpler, faster ways of doing things with less waste, and lastly, used continuous “audits” to ensure compliance with the latest best-practices (until these can be improved) and alignment between organizational policies and their vision (to prevent policy conflicts). This is in full alignment with the direction of the solution proposed by TOC today, but TOC has gone one step further to provide a simple yet powerful focusing mechanism to enable management to differentiate between the many things that can be improved and the few that must be improved now as well as a simple conflict resolution process for breaking any conflicts that block CI for the organization as a whole.

TOC’s underlying mindsets, its 5FS and TP together with the specific POOGI mechanisms defined for each of the TOC application solutions using two buffer monitoring mechanisms to focus CI efforts. The simplest tracks are where WIP is building up to identify where flow is being delayed (the constraint or CCR). The second use of buffers employs Pareto analysis to identify the major causes of red and black buffer status (buffer diagnostics), which provides an excellent focusing mechanism for the powerful time, waste, and variation reduction techniques of Lean and Six Sigma. Such a focusing mechanism should help to ensure that management can really focus properly, doing what

should be done and, more importantly, not doing what should not be done (which should save significant time, energy, and money). This approach should ensure that the organization achieves both growth and stability (from focusing scarce resources on CI of only the few parts or processes that must be improved to increase the flow of goal units).

In order for top management and other stakeholders to contribute actively and get engaged in change initiatives, they need to understand why the change is needed by visualizing the gap between the current state and the desired future state on the system constraint and identifying the consequences (on the system as a whole) of closing or not closing this gap. Only once that is thoroughly understood are stakeholders in a position to fully explore what to change (what to stop doing to address current UDEs) and to what to change (what to start doing without the risk of new UDEs), and how to cause the change using TOC's TP, and (when needed) other root cause analysis tools. The process is complete only once each of the necessary and sufficient changes have been sequenced into an implementation roadmap that will overcome likely implementation obstacles and is defined to the level of detail where it provides "actionable information."

The new TOC S&T provides a structured and systematic way of defining and communicating such an analysis or overall business strategy in a way that will create harmony (by removing the typical engines of disharmony) and, since it makes the assumptions underlying each change explicit, it also provides a practical auditing tool for each level and within each function in the organization.

In summary, remember that your focus will determine your reality. Focus on trying to improve everything or solve all problems and the possible becomes impossible. Focus on improving only the few leverage points and suddenly the impossible becomes possible.

References

- Ackoff, R. L. 2006. "Why few organizations adopt systems thinking," *Systems Research and Behavioral Science* 23:705-708.
- Al-Mashari, M., Irani, L., and Zairi, M. 2001. Holistic Business Process Reengineering: An International Empirical Survey. 34th Annual Hawaii International Conference on System Sciences, vol. 8.8022 (HICSS-34).
- Anderson, D. J. 2003. *Agile Management for Software Engineering: Applying the Theory of Constraints for Business Results*. Upper Saddle River, NJ: Prentice Hall PTR.
- Barnard A. 2001. Using TOC to Implement SAP: African Explosives Ltd. Case Study. South Africa: Saphila SAP User Conference Proceedings.
- Barnard A. 2003. New Developments and Innovations in the Theory of Constraints Thinking Processes. Paper presented at the TOCICO Conference in Cambridge, UK (October).
- Barnard A. 2007. Using a Simplified Theory of Constraints Approach to Achieve More with Less in Less Time within the Public Sector. TOCICO Conference Proceedings. Las Vegas, NV. (November 3-6).
- Barnard A. 2009. ABB Case Study Helping Achieve One-Simple-ABB with TOC in SAP. South Africa: SAPICS 2009 Conference Proceedings (June).

- Breyfogle III, F. W. 2008. *The Integrated Enterprise Excellence System: An Enhanced, Unified Approach to Balanced Scorecards, Strategic Planning, and Business Improvement*. Austin, TX: Bridgeway Books.
- Christensen, C. 1997. *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston, MA: Harvard Business School Press.
- Corbett, T. 1998. *Throughput Accounting*. Croton-on-Hudson, NY: North River Press.
- Ford, H. 1926. *Today and Tomorrow*. Garden City, NY: Doubleday. Reprinted 1988 by Productivity Inc.
- George, M. L. 2002. *Lean Six Sigma: Combining Six Sigma Quality with Lean Speed*. New York: McGraw Hill.
- Goldratt, E. M. 1990a. *What Is This Thing Called Theory of Constraints and How Should It Be Implemented?* Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 1990b. *The Haystack Syndrome*. Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 1999. *Goldratt Satellite Program, Session 8: Strategy & Tactics*. (Video series: 8 DVDs) Broadcast from Brummen, The Netherlands: Goldratt Satellite Program.
- Goldratt, E. M. 2008a. *The Choice*. Great Barrington, MA: North River Press.
- Goldratt, E. M. 2008b. *The Goldratt Webcast Program on Project Management: Sessions 1-5*. (Video series: 5

sessions) Roelofarendsveen, The Netherlands: Goldratt Marketing Group.

Goldratt, E. M. 2009. "Standing on the Shoulders of Giants." *The Manufacturer*. June. http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants. (accesses February 4, 2010).

Goldratt, E. M., and Cox, J. 1984. *The Goal: Excellence in Manufacturing*. Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. and Cox, J. 1986. *The Goal: A Process of Ongoing Improvement*. rev. ed., Crotonon-Hudson, NY: North River Press.

Goldratt, E. M., Goldratt, R., and Abramov, E. 2002. *Strategy & Tactics*. Unpublished.

Imai, M. 1986, *Kaizen: The Key to Japan's Competitive Success*. New York: McGraw-Hill/Irwin.

Kaplan, R. S. and Norton, D. P. 2002. *The Strategy-Focused Organization: How Balanced Scorecard Companies Thrive in the New Environment*. Boston, MA: Harvard Business School Press.

Knight, A. 2003. Making TOC the Main Way of Managing the Health System. TOCICO Conference, Cambridge, UK.

Kotter, J. P. 1996 *Leading Change*, Cambridge, MA; Harvard Business School Press

Mabin V. J. and Balderstone, S. 1999. *The World of Theory of Constraints: A Review of the International*

Literature. Boca Raton, FL: St. Lucie Press.

May, G. and Adams, G. 2007. A Case Study: The Change and Challenge in Engine Maintenance at Delta Airlines. TOCICO International Conference, Las Vegas, NV: TOCICO, November 3-6.

McDermott, I. and O'Connor J. 1997. *The Art of Systems Thinking: Essential Skills for Creativity and Problem Solving*. Hammersmith, London. Thorsons Publishing.

Ohno, T. 1978. *Toyota Seisan Hoshiki*, Tokyo, Diamond Publishing, 1st English translation 1988, *Toyota Production System: Beyond Large-Scale Production*, NY, Productivity Press.

Ormerod, P. 2006. *Why Most Things Fail: Evolution, Extinction and Economics*. New York: Pantheon Publishing.

Pande, P. S., Nueman R. P., Cavanagh, R. R. 2000, *The Six Sigma Way: How GE, Motorola, and Other Top Companies are Honing Their Performance*. New York, NY: McGraw-Hill.

Pink, D. 2007. *Drive: The Surprising Truth about What Motivates Us*. New York: Penguin Group.

Pirasteh, R. 2007. TLS (Theory of Constraints Lean Six Sigma) Continuous Improvement Trio—Is It Not Time to Think Differently? TOCICO International Conference, Las Vegas, NV: TOCICO, November 3-6.

Senge, P. M. 1990. *The Fifth Discipline: The Arts & Practice of the Learning Organization*. New York: Doubleday Publishing.

Shimokawa K., Fujimoto T. 2009. *The Birth of Lean: Conversations with Taiichi Ohno, Eiji Toyoda and Other Figures Who shaped Toyota Management*. Cambridge, MA: Lean Enterprise Institute Inc.

Tucker, R. B. 2002. *Driving Growth Through Innovation: How Leading Firms Are Transforming Their Futures*. San Francisco, CA: Berrett-Koehler Publishers.

Zephro, C. 2004. Integrating the TOC thinking process & Six Sigma at Seagate Technologies. TOCICO International Conference Proceedings, Miami, FL (October).

About the Author

Dr. Alan Barnard is the CEO of Goldratt Research Labs, a director of Goldratt Group (Africa), Chairman of Realization Africa and acts as project auditor on a number of large TOC projects around the world. One of Alan's passions in life is to apply the same techniques of identifying and unlocking inherent potential within businesses to individuals (he is the lead facilitators of the annual Odyssey Program) as well as to not-for-profit organizations and governments to help them achieve more with less in less time. Alan is also past chairman of TOCICO (2003-2005), Past President of SAPICS (1998 -2002), Chairman of the TOC Odyssey Institute and has been a serving judge in the National Logistics Achiever awards for Southern Africa for the last 7 years. Alan obtained his BSc Industrial Engineering degree (Cum Laude) in 1991 and completed his MSc and PhD in the management of technology and information at the Da Vinci Institute of Technology in 2009. The title of his PhD thesis was "*How to identify and unlock inherent potential within organizations*

and individuals using the system's approach of Theory of Constraints".

Appendix A—Continuous Improvement Opportunity Templates

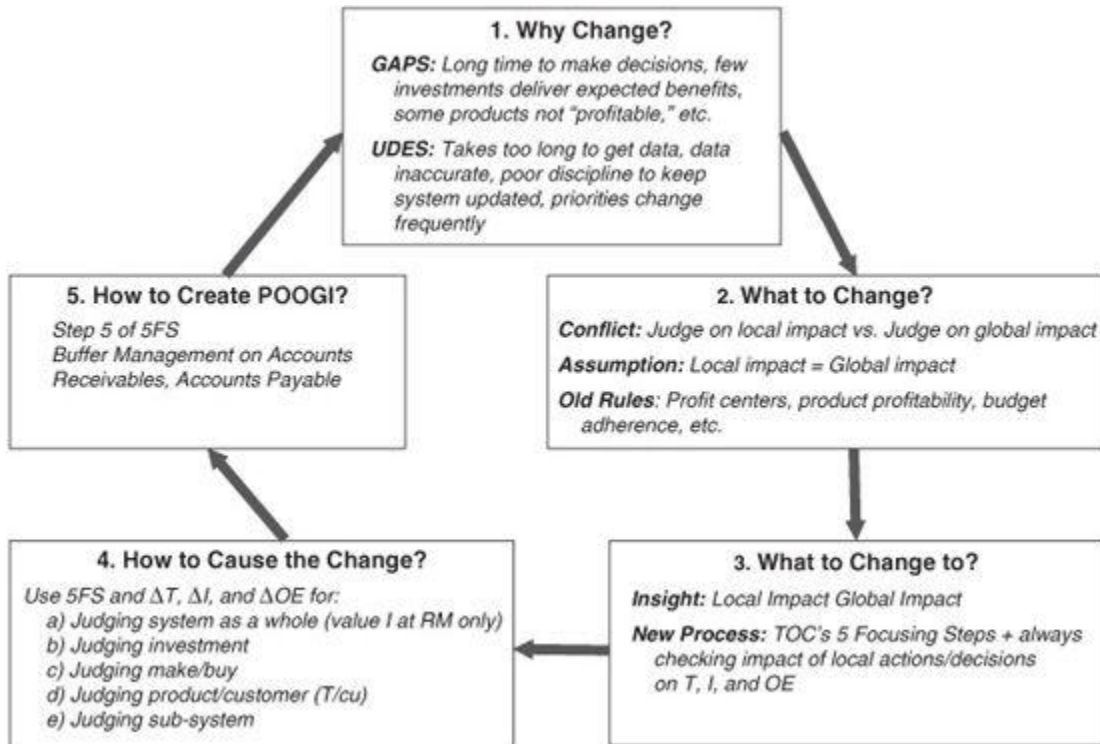


FIGURE 15-21 Applying the five questions: Managing finance the TOC way.

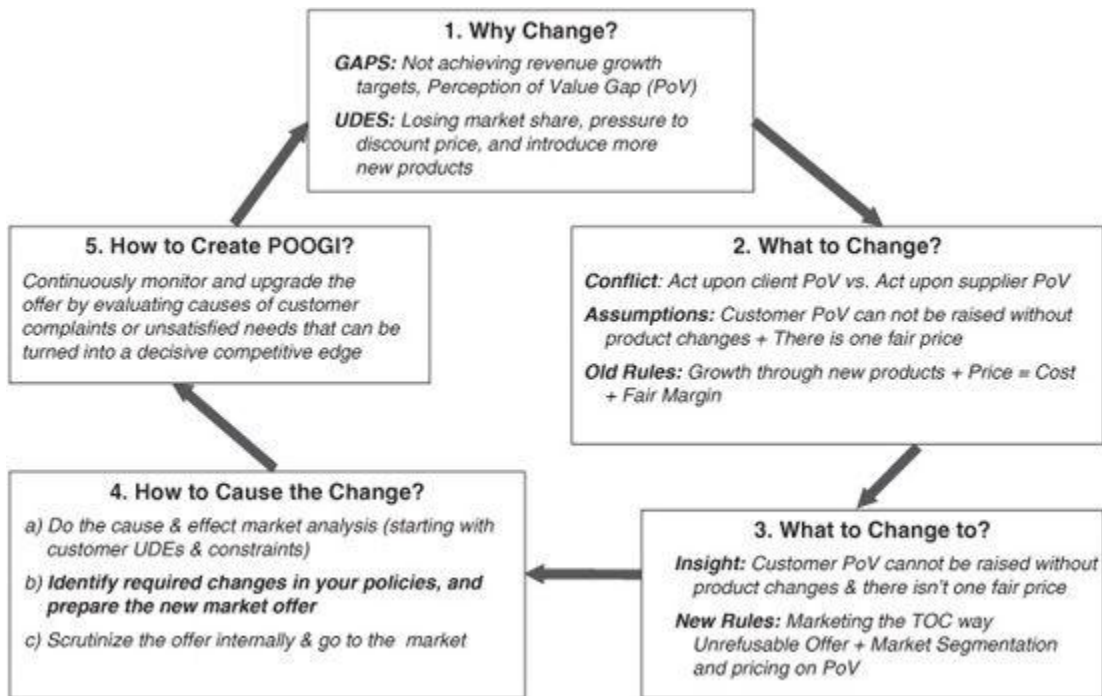


FIGURE 15-22 Applying the five questions: Managing marketing the TOC way.

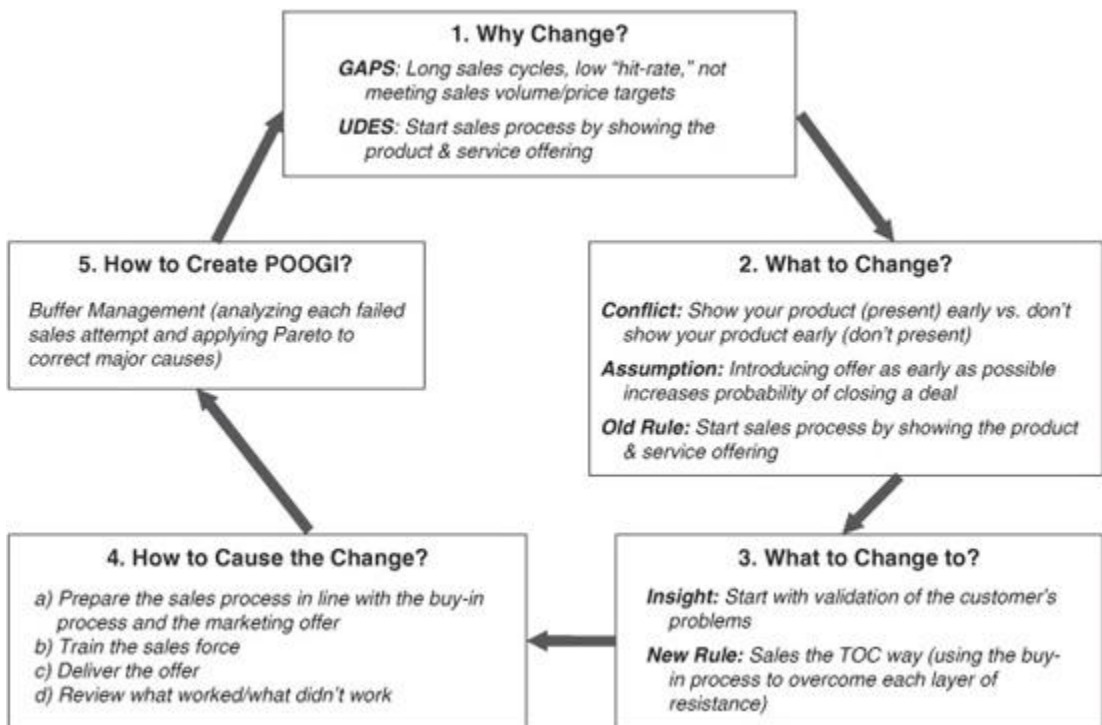


FIGURE 15-23 Applying the five questions: Managing sales the TOC way.

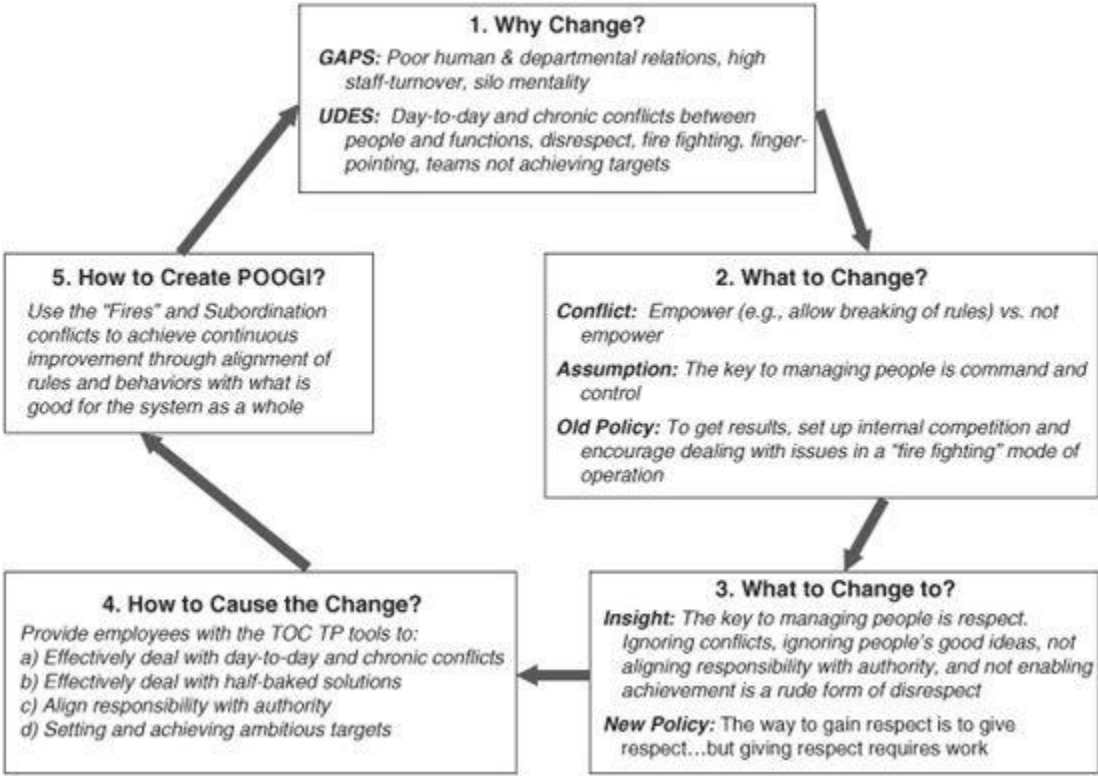


FIGURE 15-24 Applying the five questions: Managing people the TOC way.

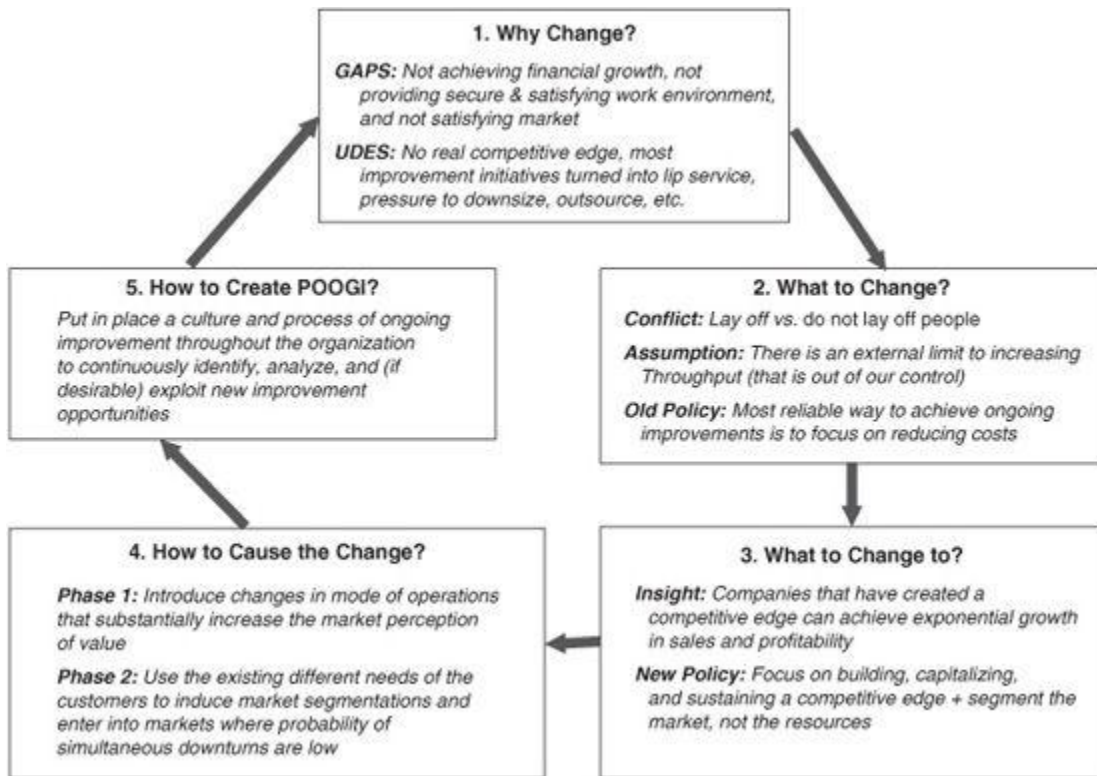


FIGURE 15-25 Applying the five questions: Business strategy the TOC way.

CHAPTER 16

Holistic TOC Implementation Case Studies

Lessons Learned from the Public and Private Sector

Dr. Alan Barnard and Raimond E. Immelman

Introduction

This chapter provides a historical perspective on the need for a holistic approach to implementing *Theory of Constraints* (TOC) as well as sharing the experiences and insights of two TOC experts (Dr. Alan Barnard and Ray Immelman), each with many TOC implementations behind them, through two case studies of holistic TOC implementations—one from the public sector (solid waste management in African cities) and one from the private sector (First Solar Inc.).

Historical Perspective to Holistic TOC Implementations

Most people are introduced to TOC through reading *The Goal* (Goldratt, 1984). A common question by readers after reading it is, “How can we replicate in our organization Alex Rogo’s achievements described in *The Goal*?”

Since *The Goal* was first published in 1984, the application and implementation of TOC have resulted in thousands of success stories from almost every imaginable industry and

type of organization. The facts related to most TOC implementations are very impressive.

- Almost anyone who tries implementing TOC achieves results.
- Results are normally very impressive, frequently above what was assumed to be possible.
- Results are normally achieved within a very short period.
- Most of the time, almost no real additional costs or investments are needed.

Copyright © 2010 by Dr. Alan Barnard and Raimond E. Immelman.

Considering these assertions, considering the pressures most organizations are under to do more in less time with the same or less resources, and considering that many large organizations today already have one or more success stories of TOC within their own organization or at least within their industry, one would expect that more organizations would be attempting to implement TOC and especially attempt to implement TOC in a holistic way.

Unfortunately, most of the past successes are from implementing TOC in only one or a few parts of the organization. Such “local” implementations always carry with them significant risks and lost opportunities (as shown by a few cases of TOC success stories that did not stand the test of time¹), which include:

1. Implementing TOC in an area (non-constraint) where the improved level of performance is already at a “good enough” level and therefore an improvement in

this area, will not translate into more goal units for the organization.

2. Improving a constraint to the level where the constraint moves to another link in the chain, and then being forced to cut “excess” capacity in the TOC link most likely killing any further continuous improvement initiatives.
3. Not capitalizing on improved performance of an area by turning it into a competitive advantage that can get customers to pay more or buy more (just a 1% increase in sales volume or average selling price from such a competitive advantage can result in a 10% or 20% increase in net profit)
4. Not reapplying TOC’s *Five Focusing Steps* (5FS) when the constraint moves or not elevating non-constraints to keep the constraint in the same place.

Goldratt himself, as far back as the late 1990s, started sharing his concerns about the fact that in his view, most TOC implementations are not holistic and reminded TOC practitioners and consultants that the real goal was not simply to make more money or achieve more goal units, but as per the subtitle of *The Goal*, the real goal was to put organizations on “*a process of ongoing improvement*” (POOGI) and ultimately turn them into “ever-flourishing organizations (the desired outcome of a Holistic TOC Implementation).”

However, how do you put a whole organization (not just Operations, Distribution, Engineering, or Sales) on a POOGI and do¹ it in a way where the implementation of all the win:win:win changes is synchronized to achieve both ongoing exponential performance growth, value, and improved stability/security for all stakeholders?²

As it turns out, answering this broad question was a bigger challenge than anyone had ever imagined. No wonder that in order to develop the first answer to it, it took Goldratt, with the assistance of hundreds of dedicated TOC experts and practitioners, an additional 15 years of work from the first publication of *The Goal*.

The Goldratt Satellite Program

In December 1999, Goldratt invited managers around the world to join him for a series of eight satellite broadcasts to share his insights on overcoming this challenge. In these sessions, Goldratt (1999) promised to share not only the essence of each of the TOC solutions, but also what he believes are the main obstacles standing in the way of a holistic implementation of these solutions.

Goldratt claimed that the obstacles have little, if any, relation to the type of industry to which an organization belongs, to its economic environment, or even to its size. Most, if not all obstacles have one thing in common: they are all human-related, psychological obstacles. Goldratt believed that identifying these obstacles and designing a process to overcome each of them successfully was the key to implementing TOC in a holistic way.

In the eight sessions, each session dealing with a different part of the organization and therefore a different TOC solution, Goldratt presented the answers to three questions. The questions were: *What to Change*, *to What to Change*, and *How to Cause the Change*. The sessions included the TOC way of managing Operations, Finance and Measurements, Product Development and Engineering, Distribution, Marketing, Sales and Achieving Buy-In, Managing People, and finally the TOC way for developing a constraint Business Strategy to synchronize the implementation of all the TOC applications.

For those organizations that participated in this global initiative, Goldratt shared his own insights and the experiences of organizations that attempted to implement TOC in a holistic way through a series of 14 letters titled “POOGI Forum Letters” (Goldratt, 1999b). In Letter 12, Goldratt directly addressed the challenge of “How to Implement TOC Holistically?” and referred again to the “X-Y Syndrome” (covered in Letter 8) which describes the dilemma of “X” (the department that used TOC to improve) when the constraint moves to another department (“Y”).

The X-Y Syndrome of Local TOC Implementations

Goldratt argued in Letter 8 that organizations faced a major dilemma in embarking on a holistic approach. In order to improve the performance of an organization continuously, the management of that organization must move the organization on a holistic approach, which requires them to start with a confined TOC application. At the same time, to improve the performance of an organization continuously, the management must not start another “short-lived” program, which requires them not to start with a confined TOC application.

Goldratt proposed that the way to break the conflict was to challenge the assumption that “the only way to start a holistic approach is to start with a local project.” However, where should we start?

- In the section which is the most receptive?
- In the section which is the most representative?
- In the section which is the constraint?

Goldratt warned that none of the above answers are valid due to the *X-Y Syndrome*. He challenged managers to think about what happens to a part or link that is managed according to a holistic approach (X) while the rest of the organization (Y) is managed conventionally. (See [Fig. 16-1](#).)

It is clear that even if X was the constraint initially, soon their improvements will be blocked by another upstream or downstream link (Y) since the significant improvement of “X” (e.g., Operations) will create major pressure on “Y” (e.g., Procurement, Engineering, Sales, or even Marketing). When Y fails to improve its performance, “X” will start feeling frustrated and will start fighting the entire system or chain. In most cases, this will be a losing battle.

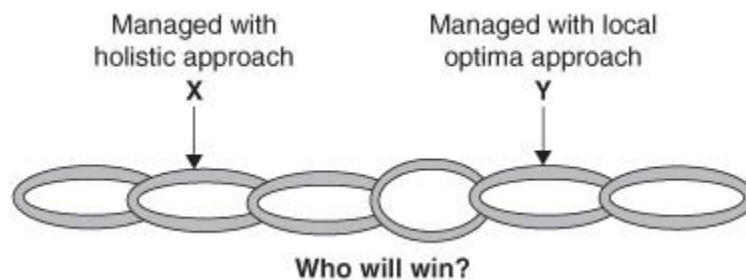


FIGURE 16-1 The X-Y Syndrome—who will win?

Since reality has shown that although possible, it is unlikely that X’s improved performance from applying the TOC/holistic approach will inspire other upstream and downstream links to follow suit, the question is where or maybe how to start with the implementation of a holistic approach such as TOC?

Goldratt proposed that rather than start with a local implementation or pilot (starting a holistic approach with a local implementation is an oxymoron), why not start by involving all key stakeholders to devise a holistic strategy for the organization?

The “4 × 4”—First Attempt at a Process to Launch a Holistic TOC Implementation

To implement TOC holistically throughout an organization requires two important paradigm shifts. The first relates to the inherent potential that can be unlocked in most organizations and within each function of the organization simply by changing the rules from the traditional way or *Cost World*³ to the TOC way or Throughout World.

The second paradigm relates to the importance of people and the fact that it is possible, with the right approach, to turn resistance to change into an active contribution to change, as long as the change makes sense and as long as it is based on a win-win-win solution.

Goldratt recommended that a holistic approach to implementation should start (if possible) with exposing key stakeholders to these two necessary paradigm shifts through understanding the TOC solutions for each link and then guide stakeholders to use the insights gained to devise a winning strategy for their organization, backed up by prudent tactics all laid out in a detailed action plan that must be achieved with the consensus of all stakeholders.

Moving the organization on a holistic approach in this way, Goldratt argued, could potentially break the conflict of where to start, since the first “local project” will simply be the first step to implementing a holistic strategy for the organization.

However, Goldratt warned that to reach “true consensus” it is not enough that each top manager agrees on the outcome, or that each top manager feels that his or her concerns (at least those that are important and urgent) have been adequately addressed. To ensure that the resulting action plan will be aggressively implemented, each top manager should rightfully feel that his or her contribution is vital and that the outcome is his or her baby; that is, all top managers take full ownership and responsibility to make this work.

However, Goldratt warned that a number of obstacles could block such an objective:

- The strategic direction concentrates only on one side of the existing conflict (local versus global optima or short-versus long-term).
- The strategic direction is nothing but polishing an existing compromise such as focusing on further reducing process cycle times when the major cause of delays is batching of orders.
- The suggested tactics are based on prevailing erroneous assumptions such as improving efficiencies is the best way to reduce cost per unit.

To overcome these obstacles and achieve the full objective of developing a holistic strategy for the organization, Goldratt proposed a new process—the 4 × 4.

The first “4” refers to 4 days where management, through the assistance of a TOC-expert facilitator, would gain a deep understanding of all the major cause-and-effect relationships governing the organization. They would be exposed to the prevailing erroneous assumptions within each of the main functions of the organization and become familiar with the common sense solutions (of TOC) stemming from a holistic approach.

Goldratt also warned against obstacles that could derail achievement of the outcomes for the first 4-day session. He cautioned against:

- The tendency to jump ahead (working on a solution when there is not yet agreement on the problem).
- The dominant person (who makes it difficult for others to contribute).

- Lack of discipline in formal logic (accepting any reservation or idea without challenging it).

To overcome these obstacles and to ensure there is a true consensus on the holistic strategy that needs to be implemented, Goldratt proposed the second “4”—4 days invested into doing a full TOC *Thinking Processes* (TP) analysis on the organization to achieve consensus on What to Change (Day 1), to What to Change (Day 2), and How to Cause the Change (Day 3), with Day 4 as a buffer.

Goldratt proposed the following steps (following the three-cloud approach as presented in Chapters 24 and 25):

To reach true consensus on “***What to Change?***”

1. Each individual contributes the biggest *undesirable effects* (UDEs) that exist in his or her area of responsibility with respect to achieving the organization’s goal (or closing the gap).
2. Each individual presents his or her conflict cloud that blocks them from removing their UDEs.
3. The group reaches the *generic or core conflict cloud* (the deepest conflict blocking growth and stability for the organization as a whole of which all the other UDE conflicts are just examples of) for their organization.

To reach true consensus on “***to What to Change?***”

4. The group exposes the erroneous assumptions underlying the generic conflict cloud.
5. The group identifies the direction of the solution that could break the generic conflict cloud.

6. Each individual applies the direction of the solution to his or her UDE conflict cloud to generate specific injections (to remove their specific UDEs and to prevent possible negatives on any stakeholder related to the new solution).
7. The group adds the missing injections to the main direction.

To reach true consensus on “**How to Cause the Change?**”

8. Each individual contributes obstacles that might block the implementation of the injections.
9. Each individual turns his or her obstacles into intermediate objectives.
10. The group builds the roadmap showing the dependencies between all the injections and intermediate objectives.
11. The group translates the roadmap into an implementation plan with each action assigned to a specific manager, together with estimated duration, etc., from which a Critical Chain plan can be built.

A significant number of organizations followed Goldratt’s advice and started their journey to implement the TOC holistic approach with the 4 × 4 process. The 4 × 4 process seemed to work where other approaches failed because it helped overcome some obstacles that other approaches could not (Kendall, 2001). However, reports from TOC champions and supporting TOC experts made it clear that there were still many obstacles blocking organizations from implementing the necessary and sufficient changes to put their organizations on what Goldratt called “the red and green curve” of exponential growth (red curve) and improved stability (green curve).

One of the major remaining obstacles was how to get the owners/executives of the company to become the internal champions for the TOC implementation - a necessary condition for both a holistic and sustainable implementation of TOC.

The Viable Vision Initiative

In November 2002, Goldratt wrote in a letter, "When I do an analysis of a company I am somewhat satisfied only when I clearly see how it is possible to bring the company to have, in less than four years, net profit equal to its current total sales." In 2003, Goldratt, rather than keeping the vision to himself, shared it with company owners, executives and senior managers around the globe, not just the vision, but also the reasons why he believed this incredible vision was viable for most organizations. Many company owners and executives responded positively to Goldratt's invitation to partner with the Goldratt Group and other TOC providers to validate if such an incredible vision was possible for their organizations.

This provided an excellent research opportunity to develop further the process to help organizations develop a business strategy that is based on building, capitalizing, and sustaining a decisive competitive edge by satisfying customers' important needs to the extent that no significant competitor can and also for learning how to really turn company owners and executives into the internal champions to personally drive the execution of the business strategy.

Since 2003, many lessons have been learned, not only how to increase the probability of achieving and sustaining the desired growth and stability for organizations, but also what changes in the process and solution design should be made when applied to different types of organizations. At the same time, we know that we are in no way there yet.

Every time we experience an inconsistency between what we expect and what happens in reality, it provides us with a valuable opportunity to develop and test a new hypothesis of what it will take to really put an organization on the journey toward an ever-flourishing organization.

Using TOC's Strategy and Tactic Tree to Guide Holistic Implementations

One of the important lessons learned by studying what worked well and what did not work well within the Viable Vision projects specifically and within other holistic TOC implementations was the need for a mechanism to define, validate, and communicate all the necessary and sufficient changes needed for achieving growth and stability as well as the sequence in which these have to be implemented. Goldratt realized this challenge could be solved with the *Strategy and Tactic (S&T) tree*, a TP tool he invented in the 1980s but formally documented in 2002 (Goldratt, Goldratt, and Abramov, 2002).

In 2006, Goldratt started investing time into developing generic S&T trees for the types of organizations that can turn a TOC logistical implementation into a decisive competitive edge. The intent of these S&Ts was not only to provide a mechanism to capture all the latest “how-to” knowledge related to each of the TOC applications, but also to prevent the two most common mistakes in holistic implementations (Ackoff, 2006): Errors of Commission (doing what should not be done) and Errors of Omission (not doing what should be done).⁴

However, is there a “one-recipe-fits-all” or simple roadmap that organizations can follow when considering implementing TOC in a holistic way, especially considering the differences between types of organizations and differences between organizations from the public and private sectors?

Catering to Differences within the Private and Public Sector

There are many differences between private and public sector organizations, but they share the fact that there is always a limiting factor—a system constraint—to achieving more goal units and that success is defined by both survival (stability) and growth. The decisions or actions needed to satisfy both the requirements for achieving growth and stability frequently result in conflicts within these organizations. Conflicts such as implementing changes or making decisions to satisfy the short-term needs of stakeholders but which could jeopardize long-term needs for the same or other stakeholders, or implementing changes or making decisions to improve one part of the organization that could result in the performance of other parts (or even the organization as a whole) being compromised. Probably the most important commonality is that both types of organizations, now more than any other time in history, are under incredible pressure to find innovative ways to “achieve more (goal units) with less or the same (resources) in less time.”

For private sector organizations, this challenge manifests in the continuous pressure to close the gap between actual and expected short- and long-term returns for shareholders. For public sector organizations, the challenge manifests itself in the ongoing pressure to close the large and frequently growing gap between the deteriorating levels of service delivery and infrastructure and a growing demand for such services in the areas of health, safety, education, energy, and telecommunications—especially in the developing countries around the world.

It is these gaps between the expectations of stakeholders and the current performance of an organization or system that can result in vicious cycles of over- and under-reactions or that can be the catalyst for challenging the status quo

and achieving sustainable improvement. These gaps serve as a constant reminder that there must be a simpler, faster, better, and more reliable way to identify where and how to improve our organizations.

However, the realization that a large (and potentially growing) gap exists and that scarce resources should be focused on closing this gap is not sufficient. Management must know what changes are needed to close these gaps and, as importantly, how to remove obstacles and excuses for not taking the necessary and sufficient actions to turn good ideas into sustainable results.

The next section provides the details and lessons learned of how to do a holistic TOC implementation within the public sector (the story of a sustainable cities program for Africa) and the following section provides the details and lessons learned on how to do a holistic TOC implementation within the private sector (the story of First Solar).

Holistic Implementation of TOC in the Public Sector

Over the past 25 years, TOC has helped thousands of for-profit organizations⁵ improve their performance and decision making. However, despite TOC's universality to help identify and unlock inherent potential within any goal-orientated system, few *not-for-profit* (NFP) organizations have attempted to apply TOC and even fewer have attempted a holistic TOC implementation.⁶

Typical reasons include a perception that TOC is probably too complex or sophisticated for their organization, or that since many of these NFP organizations do not have clear goal statements such as "Make more money now as well as in the future," TOC will probably not work for them. In addition, it is claimed that for any NFP to reach consensus that TOC should be implemented on a holistic basis or even just evaluated for implementation is an order of magnitude

more difficult than in for-profit organizations due to the large number of stakeholders involved.

In this case study, we want to share the experience and insights gained on how a new simplified TOC holistic analysis, consensus building, and active contribution approach can help achieve the very ambitious target of “doing more with the same or less in less time” even in NFP organizations. This new simplified win-win-win approach is based on combining TOC’s enabling paradigms (discussed later) with its Five Focusing Steps (5FS) and Thinking Processes (TP), as well as the simple planning and execution rules recommended by the logistical solutions of TOC, to help all stakeholders identify and unlock inherent potential within the “system” for which an NFP organization is responsible.

Background

In January 2007, Mr. Michael Funcke-Bartz of InWEnt-Capacity Building International, Germany, in the framework of the German Development Cooperation requested the assistance of two TOC experts (Dr. Alan Barnard, CEO of Goldratt Research Labs and Professor Antoine van Gelder, Head of the Department of Internal Medicine University of Pretoria) with the objective to test whether a simplified TOC constraint analysis and strategy development approach (Barnard, 2003), could be used to help cities close the growing gap between demand and supply, especially for basic services. By 2007, InWEnt already had a few successful TOC-based projects from the public (improving management capacities in water utilities) and the private sectors (strengthening of small- and medium-sized enterprises) in developing countries (Funcke-Bartz, 2006) but indicated they wanted to test the impact of a more

holistic approach to using TOC to supplement and potentially focus their own capacity-building efforts.

InWEnt suggested that this could be done with African cities from UN-HABITAT's Sustainable Cities Programme, which had applied for capacity-building assistance in the field of municipal solid waste management, because of its potential contribution to city sustainability and poverty alleviation.

To ensure the best possible start to such an initiative, InWEnt required these cities, to organize a one-week Strategy Workshop with representatives of all-important stakeholders such as national and local governments, public and private service providers, the community and academic institutions.

Invited stakeholders were informed that the objective of the workshop was first to work together to develop a common understanding around the cause-effect relationships between the various challenges faced by each of the stakeholders in dealing with the causes and consequences of the large GAP between the amount of waste created versus the amount of waste collected on a daily basis. Second, the aim was to agree on which critical few changes and contributions by each stakeholder, could help overcome capacity and policy constraints to close this large and growing gap now and in the future.

This opportunity not only provided a way to test whether a holistic TOC constraint analysis and strategy formulation workshop with all the stakeholders was the best possible kickoff within the public sector for a holistic TOC implementation, but also to validate whether (even for NFP organizations and their stakeholders) that TOC could be used to help *achieve more with the same or less resources in less time*. This could be done by helping stakeholders identify and challenge assumptions that are potentially limiting their ability to see and/or unlock inherent potential within themselves, their organizations or even their cities.

Such *limiting assumptions* especially if held by key stakeholders in any system, can jeopardize the ability of that system to achieve the desired or required growth and stability.

Four African cities were selected for a TOC-based capacity building process starting with a Constraint Analysis Strategy Workshop focused on solid waste management:

1. City A—population of over 3 million people
2. City B—population of over ½ million people
3. City C—population of over 2 ½ million people
4. City D—population of over 2 ½ million people

Designing the Five-Day TOC Workshop and Implementation Process

Complicating Factors within the Public Sector

Prior to the start of this initiative, the TOC expert team members were warned that the public sector is much more “complicated” than private sector organizations. Those knowledgeable in this arena warned that normally there are no clear goals and the system is full of bureaucrats which Dr. Russell Ackoff (2006) defines as someone with the power to say “No” but not the power to say “Yes”. Also, that frequently there is a high level of distrust among the groups, higher resistance to change due to awareness that the wrong policies or decisions could have catastrophic consequences, many and sometimes conflicting objectives (which makes “focusing” difficult), and little tolerance for “business principles and best-practices,” especially for a business method called *Theory of Constraints*.

The team was also informed that although there is excellent work being done by the many international donor and capacity-building organizations within developing countries, many of these initiatives struggle to resolve the underlying conflicts and to really secure the full consensus, contribution, and commitment from all stakeholders to ensure that initiatives do deliver the desired results in a sustainable way.

The next step in the design process was to use the TOC Change Management Questions to provide the agenda for the five days and to design a process around these five questions (stated below) to achieve the desired outcomes of the workshop in a way that would overcome both generic and specific obstacles in achieving consensus on where to focus scarce resources.

Simplified TOC Analysis and Implementation Roadmap

The proposal to InWEnt by the TOC Expert process design team was that the best possible start (Barnard, 2008) would be a 5-Day Constraint Analysis Strategy Workshop. The process would follow a simplified and improved version of the traditional TOC Analysis Roadmap (which only covered What to Change, to What to Change, and How to Cause the Change) and which should include the following steps:

Step 0: Introduction to TOC's Systems Approach, Processes, and Paradigms

Step 1: Consensus on *Why Change?*

Step 2: Consensus on *What to Change?*

Step 3: Consensus on *to What to Change?*

Step 4: Consensus on *How to Cause the Change?*

Step 5: Consensus on *How to Measure the Change and Achieve Ongoing Improvement?*

This roadmap, also serves as an agenda for the five days.

Day 1 aims to cover Steps 0 and 1—to get agreement on the new systems approach (transition from *limiting to enabling paradigms*) as well as to answer “Why Change?” for the system and its stakeholders being analyzed by identifying the gap in goal units, the consequences of not closing this gap, and what makes it difficult for each stakeholder to contribute to closing the gap (the UDEs).

Day 2 aims to cover Step 2 by helping stakeholders answer “What to Change?”—identifying the core conflicts for each stakeholder and the underlying assumptions and associated “local optima” rules and measurements that must be challenged to remove the UDEs (to close the gap).

Day 3 is dedicated to Step 3, answering “to What to Change?”—identifying the new win: win contributions, rules or measurements (the injections) that will break the core conflicts, remove the UDEs, and close the gap without creating new UDEs.

Day 4 is focused on completing Step 4 by answering “How to Cause the Change?”—identifying the possible risks (such as the injections that are not detailed enough to provide actionable information, the injections that have potential negatives or could be blocked by implementation obstacles) and how to overcome these, sequenced in an implementation roadmap.

Day 5 is focused on Step 5 to answer “How to Measure the Change and Achieve Ongoing Improvement?”—identifying the prime measurements for each stakeholder’s required contribution, allocating responsibilities, and (if time permits) converting the roadmap into a buffered but aggressive Critical Chain plan and sharing basic *TOC Critical Chain Project Management* (CCPM) insights for managing the implementation as a portfolio of pipelined and buffered projects.

Proposed Changes to the Traditional TOC TP Analysis Roadmap

The original TOC TP Roadmap as taught in the TOC Thinking (“Jonah”) and “External Constraint Analysis” programs or the “4 × 4” Holistic Approach has proven to be an effective analysis and strategy development toolset. However, TOC practitioners that have used this TP roadmap know that there are a number of problems with the traditional process which focused on answering only the three change questions as first formulated by Goldratt (1990)—What to Change?, to What to Change?, and How to Cause the Change? Problems frequently experienced with the original TP Roadmap include the following (Barnard, 2003):

1. Starting a TOC analysis without sharing the fundamental TOC principles and paradigms of TOC needed to change the behavior of the various stakeholders such as the beliefs in inherent potential, inherent simplicity, win-win, “good enough,” and that people are good (but sometimes have bad assumptions that drive bad behaviors) as per [Table 16-1](#).
2. Sometimes wrongly assuming there is already agreement by all stakeholders on the need for change (i.e., starting with “What to Change” rather than the gap in goal units and constraint performance and consequences for each stakeholder if this gap is not reduced to get agreement on “Why Change?”).
3. Starting the analysis simply by asking stakeholders to list their UDEs (what bothers them) rather than by asking stakeholders to contribute a short list of those few UDEs that make it difficult for them to help close the gap in goal units and constraint performance (with respect to the goal), which provides a much more

relevant list of UDEs to start the TOC TP analysis with and effectively links the TOC's 5FS with the TP.

Improvement Challenges	Limiting Paradigms	Enabling TOC Paradigm
1. Constraints	Assume most "constraints" are inherent/ "Out-of-my Control"—blame them for the gaps and focus on "In-My-Control" issues.	Assume most "constraints" are "In-my-Control or Influence" (inherent potential)—just find ways to better exploit or elevate them.
2. Complexity	Assume complexity is inherent—simplify it by breaking the system/problem up into simpler parts and optimize each part.	Assume inherent simplicity. Find the leverage points (the important few) and focus all efforts on capitalizing on these.
3. Conflicts	Assume that for me/us to win, you/they must lose. Find compromises or even win-lose solutions.	Assume win-win is always possible, just look for it.
4. Uncertainty	Assume inherent certainty—look for "formulas" to calculate optima points for decision variables such as price, capacity, stock, etc.	Assume inherent uncertainty—find "good enough" solution and use feedback to improve and sustain decision variables.
5. Bad Choices/ Behavior	Assume some/most people are just inherently bad—find and get rid of bad people.	Assume people are inherently good—find and get rid of bad assumptions.

TABLE 16-1 Improvement Challenges and the Related Limiting versus Enabling Paradigms

4. Asking stakeholders to verbalize any conflicts associated with their UDEs rather than giving specific instructions to verbalize both the *symptomatic conflicts they experience in having to deal with their UDEs* and the *systemic conflicts of the part they are blaming for causing their UDEs* as per the *dual cloud process* (Barnard, 2003). Not verbalizing both conflicts has two major negatives. The distrust between those blaming and those being blamed will continue and it will not be possible to find the "core conflict" for each stakeholder when *systemic and symptomatic* conflict are mixed together. [Figure 16-2](#) shows the separation between the systemic and symptomatic conflicts in the dual cloud approach.

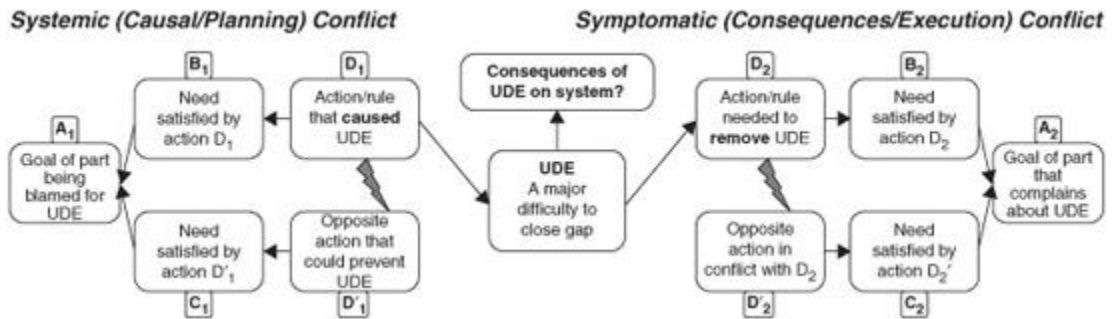


FIGURE 16-2 Systemic versus symptomatic conflicts related to an undesirable effect.

5. Assuming that there is always only one core conflict for a system, when in fact there is one core conflict for each stakeholder. This is the conflict that most contributes to blocking that stakeholder from contributing in a way that can help better exploit or elevate the system constraint.
6. Attempting to break conflicts using the traditional TOC TP approach of identifying and challenging all the many assumptions associated with the core conflict rather than focusing on finding and challenging only the one or few “conflict assumptions” using the four methods⁷ as proposed by Barnard (2003) that if removed, will “evaporate” the core conflict.
7. Not ensuring the injections needed to break the conflicts have been defined as “Actionable Information,” which means each stakeholder can clearly verbalize how they will contribute to implement the injections agreed upon to break the core conflicts blocking better exploitation, subordination, or elevation of the system constraint.
8. Not having a simple and fast mechanism to enable stakeholders to raise both types of “Yes, buts...” that

can block each stakeholder from making the contributions needed to improve the system. The first “yes, but...” relates to concerns about significant predicted UDEs that could impact one or more stakeholders if the agreed injections are implemented. The second “yes, but...” relates to concerns about significant implementation obstacles that must be overcome to implement the agreed injections. Frequently, these two types of concerns or reservations are not dealt with or confused during the analysis resulting in significant rework later or even project failure.

9. Not agreeing on what the prime measurements will be that must be implemented to validate whether the gap in goal units (e.g., between demand and supply) is really closing and the secondary measurements to validate whether each stakeholder is able to make their agreed contribution.

To ensure that organizations can get all the stakeholders together to do a TOC analysis on their “system” can get the best possible start by not making any of the mistakes listed above (especially when in most cases we will not get more than one attempt at it), a new TOC TP Roadmap was needed.

As a result, a new simplified TOC TP Roadmap was designed (see [Fig. 16-3](#)) to address each of these nine problems in a way that would rebuild trust and ensure the active contribution of all stakeholders and which can be completed (with an experienced TOC facilitator) in only five days (one day for getting agreement on each of the five change questions provided previously in Steps 1-5).

The next section provides a detailed example of how the new simplified TOC TP analysis process was used in the five-day workshop to get full agreement on the necessary and sufficient changes that must be implemented by the various

stakeholders (to help close the gap) and in what sequence they have to be done.

Detailed Case Study: Analysis on Solid Waste Management in City A

This section provides an overview of how the simplified 5-Day Constraint Analysis Process was tested in the first cities (A and B) as well as providing a summary of the outcomes achieved and generic lessons learned from the application of this process with other cities.

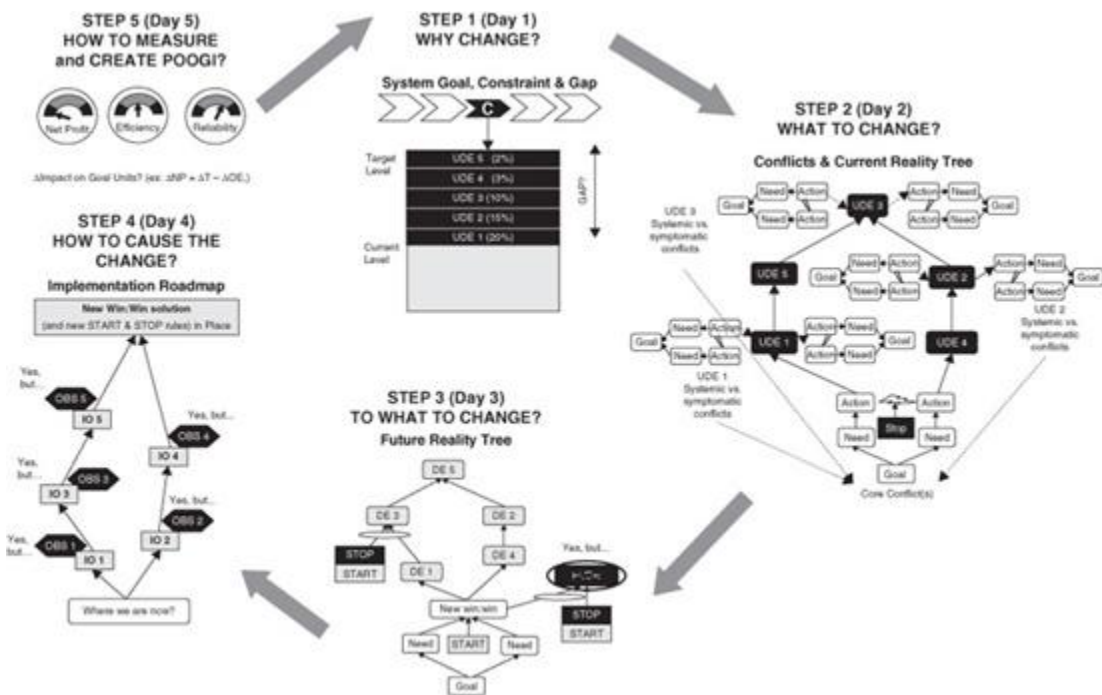


FIGURE 16-3 Barnard’s new simplified TOC analysis roadmap.



FIGURE 16-4 Typical buildup of solid waste due to service delivery gap in informal areas.

Current Reality within the Management of Solid Waste in City A

In most African cities, local governments are responsible for ensuring that garbage (solid waste) is collected, transported, and disposed of in a safe and environmentally friendly manner. Common practice is for private contractors to collect waste from private businesses and households in formal areas and for *Community-Based Enterprises (CBEs)* to collect waste from informal areas. However, often in developing countries these services do not cope with the demand (even in the formal parts of the cities) and generally do not reach the very poor because they live in areas that are difficult to access or they cannot pay for the service.

The results of this gap in service delivery can be seen everywhere in these “un-serviced” or informal areas ([Fig. 16-4](#)). Huge piles of garbage (solid waste) can be found throughout poor settlements, making poverty, persistent ill health, and low environmental quality part of everyday life.

In view of this, the City Council in conjunction with InWEnt, organized a strategic Constraint Analysis and Planning Workshop with the invitation to stakeholders listing the following objectives of this workshop:

1. Joint analysis with involved stakeholders for a better understanding of the complex relationships between the different problems related to waste management services and their causes.
2. Identification and prioritization of capacity constraints and/or policies, measurements or behaviours that prevent the improvement of *solid waste management* (SWM) in the city.
3. Agreement on the critical contributions and capacity-building strategy to cope with the identified problems.

All stakeholders representing the different points of view and interests involved in this matter were invited to attend the workshop that was facilitated by Dr. Alan Barnard and Professor Antoine van Gelder (both TOC Experts)⁸ and supported by the InWEnt team of Michael Funcke-Bartz and Maria Sagué as well as the Kenyan project coordinator Stanley Mbagathi.

Step 0—Creating the Shift to a TOC/Systems Approach Paradigm

The workshop is opened by the most senior representative of the City Council (normally the Mayor or Town Clerk) to share the objective of the workshop and why they have invited all stakeholders to contribute to the analysis and solution development (to ensure it is a winwin-win). Then a representative from InWEnt shared why they have selected TOC as the approach to do a holistic analysis on the system and why they partnered with the TOC experts to facilitate the workshop (to establish credibility).

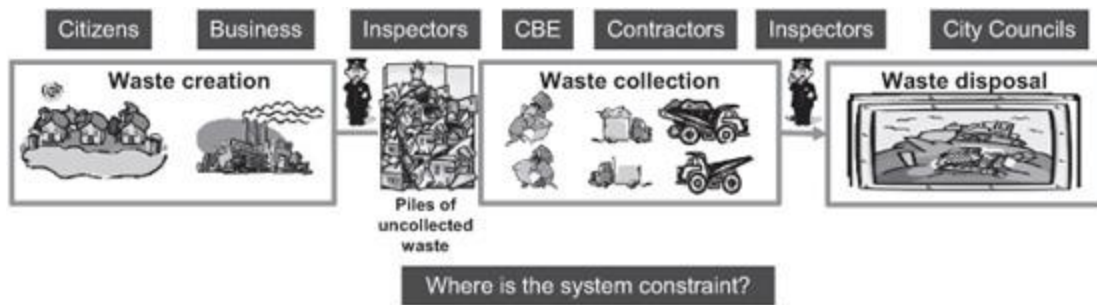


FIGURE 16-5 Example of the solid waste management chain/system.

During the morning session, the TOC facilitator introduces TOC as a simple and powerful way to overcome the five challenges listed in [Table 16-1](#), faced by all stakeholders in both private and public organizations (constraints, complexity, conflicts, uncertainty, and bad choices/behaviors). To show the magnitude of the inherent potential that can be unlocked if the right “limiting assumptions” are challenged and replaced with the right rules (that enable better exploitation of the system constraint), the TOC expert can use an interactive simulation game such as the “multitasking” game⁹ that shows how, by just changing from a rule of “multitasking” (based on an assumption that the earlier we start the earlier we finish) to a rule of “no multitasking” (based on a realization that the later we start the sooner we finish) we can do double the amount of projects in half the time with the same resources.

The simulation game also provides an opportunity to introduce TOC’s 5FS and TP analysis process of defining the goal, gaps, consequences, UDEs, core conflict, injection, and associated actionable changes in the rules, and potential negatives, and how to prevent these and implementation obstacles and how to overcome these. The introduction part ends with relevant success stories of where TOC has been applied and introduces the roadmap and desired outcomes

for the rest of the five days, inviting everyone to contribute actively. An example of the slides used for such an introduction can be downloaded at www.goldrattresearchlabs.com (in the “Downloads” section).

Step 1—Getting Agreement on the System Constraint and Why Change?

The afternoon of Day 1 starts with representatives of each of the stakeholders coming forward to draw the “system” as a chain and to identify where they believe the weakest link (system constraint) is based on where most of the “stuff” (garbage) is piling up (right before waste collection) as per [Fig. 16-5](#).

Agreement on “Why Change?” is achieved by identifying the current and future gaps (between the demand—tons of solid waste created everyday by citizens and businesses—versus the supply—tons of solid waste actually collected and disposed of) and getting each stakeholder to share their view of the negative consequences on all stakeholders if the gap is not closed soon. [Figure 16-6](#) shows an example of what the current and future gaps (based on expected growth in population of the city) looks like. Typical “consequences of not closing this gap” include disease, increased demand on already stretched health services, environmental time bombs, etc.



FIGURE 16-6 Current and future gaps of waste created versus collected in City A.

Before moving to the next question (What to Change to) stakeholders are then asked to identify the UDEs that make it difficult for them to contribute to closing the gap (e.g., either by reducing the tons of waste created or by increasing the tons of waste collected) and to share what initiatives have been put in place in the past to deal with these problems. Of course, the fact that the UDEs still exist means that these past initiatives were not successful or, at most, were partially successful. [Table 16-2](#) shows an example of the typical UDEs raised by stakeholders, what the common belief is on the root causes of these UDEs, and the traditional solutions that have not been effective at removing these UDEs.

Getting stakeholders to agree that the past “solutions” such as more budget or more education have not made a significant impact on closing the gap is an important step before moving to “What to Change to”; otherwise, there is a risk that previous assumptions of root causes and solutions will simply be tabled again.

Day 1 ends with each of the stakeholders presenting their list of UDEs and why they are “undesirable” based on their

impact on the rest of the system.

Problems (UDEs)	Root Cause	Traditional Solution
Insufficient government budget	Lack of awareness	Awareness campaigns
Residents do not want to pay		
Service is too expensive		
People dump/burn illegally	Poor/no enforcement	Stricter enforcement of bylaws
No waste bins available		
Frequent equipment breakdowns	No equipment or old equipment	Get external funding to buy/build necessary infrastructure
Cannot access all areas		
No service in some areas	Low/no subscription rate	More awareness campaigns
Actual Waste Collection = 300 to 400 Tons per day (versus Actual Waste Created = 1000 Tons per day)		

TABLE 16-2 Conventional Way to Deal with UDEs in Waste Management in African Cities

Step 2—Getting Agreement on What to Change?

Day 2 starts with the TOC expert facilitator introducing the TOC definition of a problem not only as a gap between reality and expectation, but as a set of unresolved conflicts related to closing the gap. An example of a common UDE such as high inventory in a retailer can be used to get stakeholders to experience the process of verbalizing the symptomatic conflict for the one having to deal with the UDE (e.g., Sales under pressure to reduce prices to get rid of the inventory) and verbalizing the systemic conflict for the one being blamed for the UDE (e.g., Purchasing under pressure to continue to buy large quantities to secure volume discount).

Each of the stakeholders then selects the three UDEs they think contribute most to the current gap and then work in their group to verbalize their symptomatic and systemic conflicts. Once completed, a representative of each stakeholder presents it to the whole group. We have found that this dual cloud approach is not only an effective way to

verbalize the “real problem,” but it also forces each stakeholder to look at their problem (their list of UDEs) from the view of the one they are currently blaming (i.e., they can show understanding for the conflict in which the one being blamed is stuck). This approach plays a critical part in rebuilding the trust between the stakeholders.

For City A, there were four stakeholders (City Council, CBEs, contractors, and residents). Figures 16-7 and 16-8 list the four sets of systemic conflicts and symptomatic conflicts for each of these stakeholders. Note that the [D] actions in the systemic conflicts are the actions being blamed for causing the UDEs in the first place while the [D] actions in the symptomatic conflicts are the actions the specific stakeholder feels most pressure to take to deal with their UDEs. In the case where the one being blamed for the UDE is also the one having to deal with the UDE, the systemic and symptomatic conflicts will be the same but with swapped [D] and [D’].

The process of each stakeholder group verbalizing, validating with the facilitator, presenting it to the whole group, and improving verbalization of their symptomatic and systemic conflicts based on feedback from others normally takes between a half and a full day (e.g., when there is a large number of stakeholders involved).

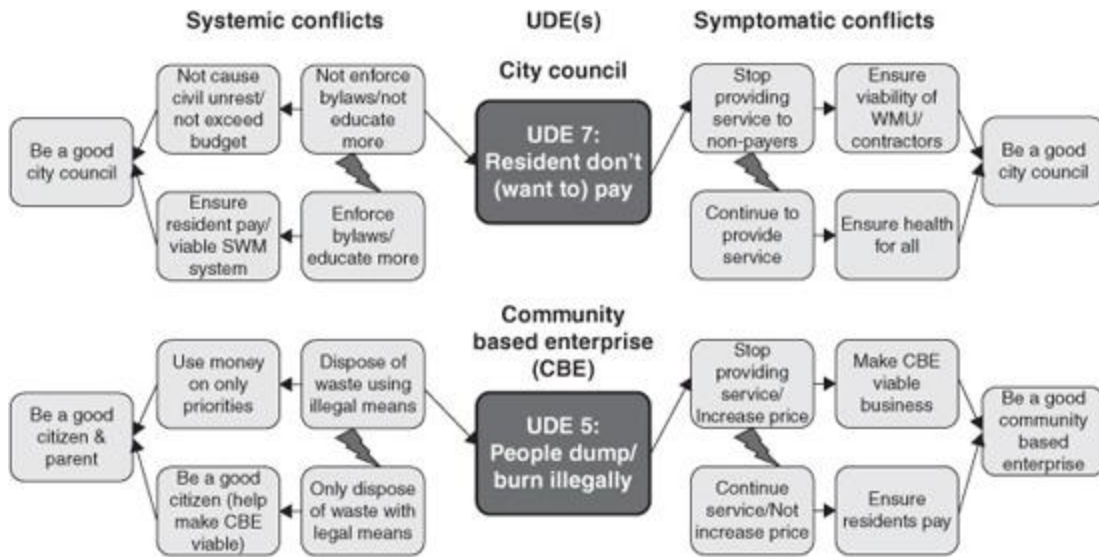


FIGURE 16-7 Systemic and symptomatic conflicts for city council and CBEs.

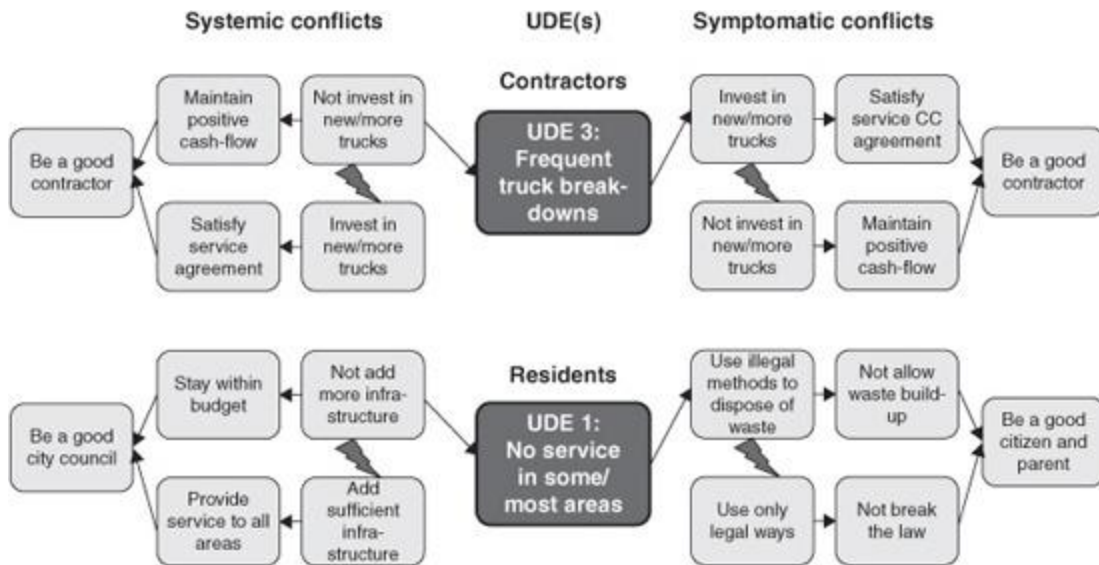


FIGURE 16-8 Systemic and symptomatic conflicts for contractors and residents.

The next step in the process is for the TOC expert facilitator to guide the identification of the core conflict for each stakeholder using the normal process of writing down all the Ds, D's, Bs, Cs, and As and asking the participants to

look for a generic pattern (i.e., “All the Ds can be summarized by *Don’t invest/spend more*”).

Tip: Starting with the Ds and D’s makes it a relatively simple process to verbalize and validate the rest of the core conflict for each stakeholder.

Normally, this convergence is quite simple (but not easy especially for novices) because typically the systematic conflict of one stakeholder is the same as the symptomatic conflict of the one they are blaming.

[Figure 16-9](#) shows an example of what the four core conflicts for each of the four stakeholders are and how they are related.

Once completed, a representative of each stakeholder presents their core conflict to the group. During these presentations, we normally receive a remarkable level of consensus on the first verbalizations, which definitely elevates the credibility of the process in the eyes of the stakeholder participants as they could, from previous days, clearly identify with the core conflicts of each stakeholder. We did experience a few situations where stakeholders suggested better verbalizations, which were immediately accepted by all and incorporated into the final versions.

Step 3—Getting Agreement on What to Change to?

In the third step (normally the third day) of the analysis, the stakeholder groups work on their core conflicts, using the four methods (Barnard, 2003) to quickly identify and invalidate only the conflict assumptions (in contrast with the traditional approach that attempts to identify and invalidate all assumptions related to the conflict). Then stakeholders define possible breakthrough win-win solutions (injections that will invalidate the conflict assumptions and therefore break the core conflict). [Figure 16-10](#) shows an example of one of the outcomes of applying Barnard’s four-method

process for breaking conflicts to the service provider's investment core conflict.

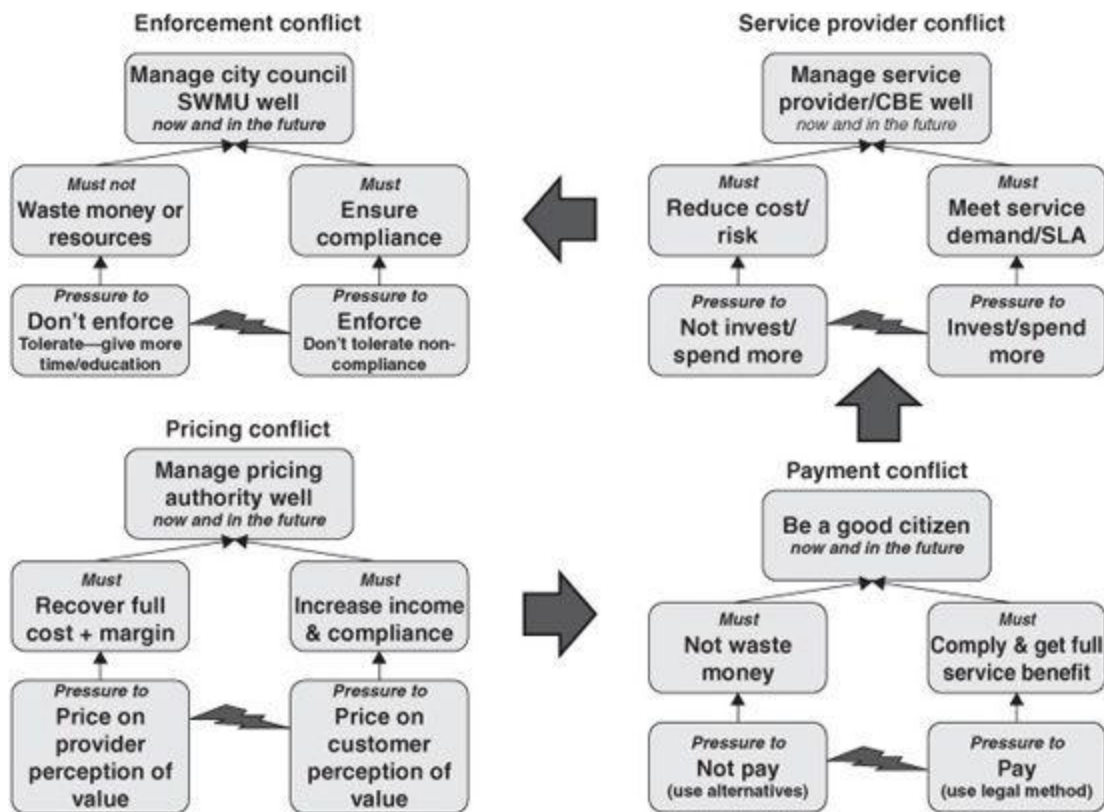


FIGURE 16-9 Four core conflicts identified for each of the four stakeholders.

The generic injections to break the four core conflicts, as identified with this method, provide the answer to the “to What to Change?” question for each stakeholder. The four generic injections for the four core conflicts are shown in [Fig. 16-11](#).

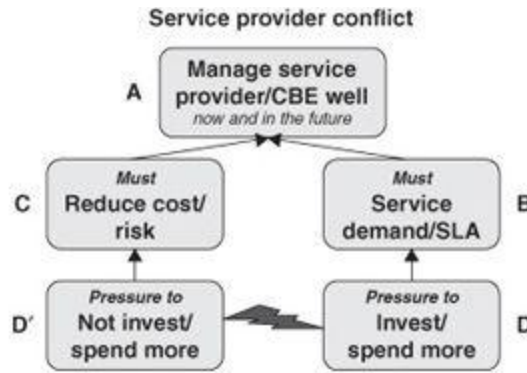
Once the injections (which provide the direction of a win-win solution) and some of the potential “How to” (viable ways to make the injections actionable) and benefits (to validate if the injection will remove the original UDEs) have been identified by each stakeholder for their core conflict, representatives of each stakeholder again present their summary of “to What to Change” to the group. This can

help with improving verbalization and suggestions of other ways to achieve the injections. This step is normally completed by the end of the third day or at the latest during the morning of the fourth day.

To prepare participants for the next step, the TOC expert facilitator then asks the group to contribute their “Yes, buts...”—the reasons normally given why something will not work or why it will be impossible to implement. Rather than discouraging such “negative views,” we encourage them as another way to contribute to ensuring the planned changes will work, will not result in any major negatives, or won’t be blocked by major implementation obstacles.

Step 4—Getting Agreement on How to Cause the Change?

Step 4 (typically done on Day 4) deals with the identification of the two types of “Yes, but...” reservations—reservations on why the new injections or solutions might hurt one or more of the stakeholders (*predicted undesirable effects* or PUDEs) or why it might be too difficult to implement (*implementation obstacles* or OBSs).



Arrow	Assumption	Injection	How to . . .	Benefit
M1 – D'B	Current capacity insufficient to meet current demand	Current capacity sufficient to meet demand	Ensure we do not commit to supply more than our capacity to deliver; and know demand and compare against total capacity in region (before investing)	Meet demand at minimum cost and risk
M2 – DC	If you spend more you increase cost and risk of business	We have a way to invest more and not increase cost and risk	We approach businessman to help prepare business plan to get low-interest loan from bank	Meet demand at minimum cost and risk
M3 – DD'	Both cannot be done at the same time	We agree when to invest and when not to and also where to invest	Use backlog as early warning indicator. If backlog starts growing, we know when to invest and invest only in "Bottleneck" resources	Meet demand at minimum cost and risk
M4 – E	Only way to increase capacity is through investment	Find way to increase capacity without increasing cost	Use TOC to help better utilize capacity we have before getting more (exploit before elevate)	Meet demand at minimum cost and risk

FIGURE 16-10 Service provider's core conflict, conflict, assumptions, and injections.

In the traditional TOC TP, the Type 1 "Yes, but..." will be identified and shown in the *Future Reality Tree* (FRT) process that uses sufficiency logic to show the cause-effect of how each of the injections should result in the desired outcomes (the *desired effects* or DEs) as well as what PUDEs these injections can cause if the injections are implemented as they were originally defined in Step 3. Part of the process to build an FRT is, therefore, to identify these PUDEs as well as the additional requirements (that would be added to the main injection) to prevent these PUDEs.

In the traditional TOC TP, the Type 2 “Yes, but...” reservations are identified in the *Prerequisite Tree* (PRT) process that uses necessity logic to show the necessary conditions for achieving the successful implementation of the main injections. To identify these necessary conditions, the TOC expert and group of stakeholders tried to identify likely obstacles (OBSs) and how to overcome these with *intermediate objectives* (IOs); these will serve as implementation milestones.

To simplify these processes without jeopardizing the completeness of the analysis, a new simplified process (done in groups) was developed to get each stakeholder to help contribute “Yes, but...” reservations and suggestions on how to prevent or overcome these. All stakeholders are asked to participate in this process with the justification that the more of the “Yes, but...” reservations we can identify and proactively prevent during the analysis phase, the less surprises we will get in the implementation and therefore the higher the probability of success.

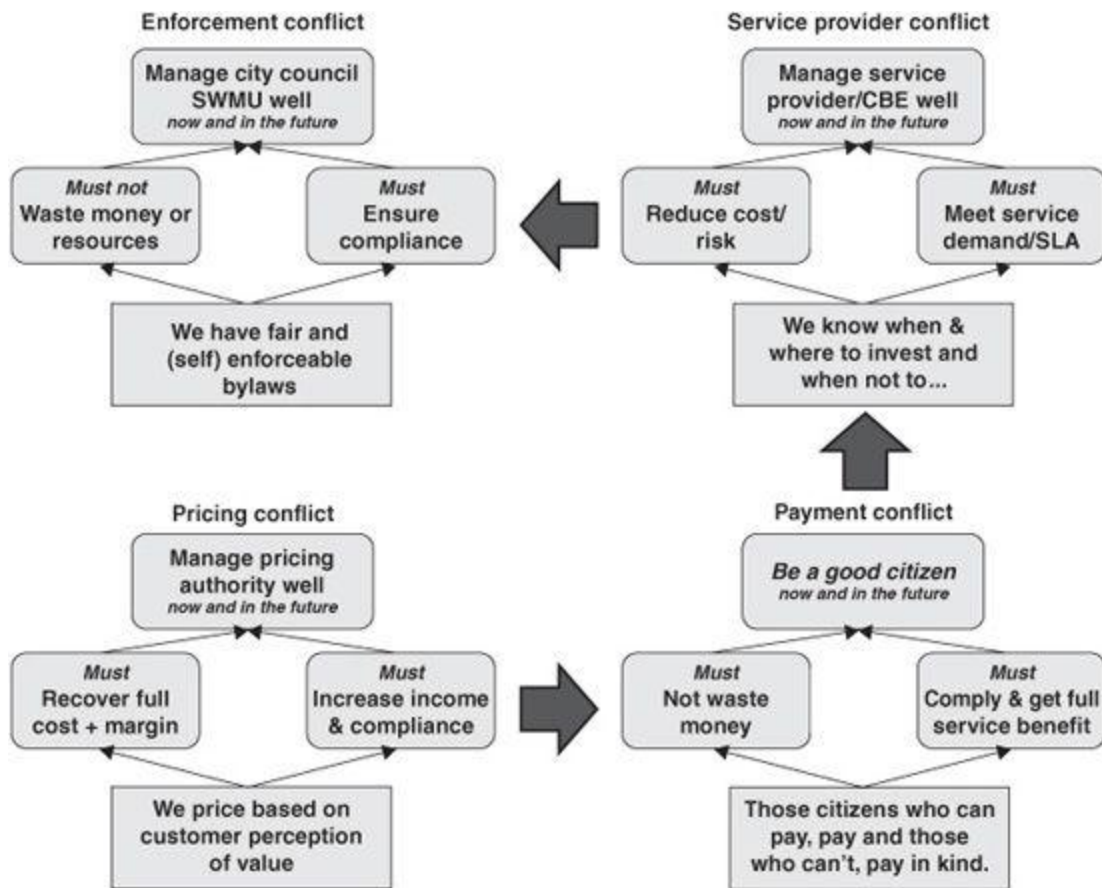


FIGURE 16-11 Four injections to break the four core conflicts for City A's SWM stakeholders.

To obtain the contribution of all stakeholders, they are divided into cross-functional groups. The members of each group are then encouraged to identify all the stakeholders that can be affected negatively or whose help will be required for the implementation. For each of these stakeholders, participants are then encouraged to contribute as many PUDEs of the planned changes and how these can be prevented (to ensure new solutions are really winwin) as well as contribute as many potential OBSs and how to overcome these (adding additional injections or additional conditions to already agreed injections). [Figure 16-12](#) shows the new simplified framework to enable

participants to identify and overcome both of the “Yes, but...” reservations in a single step.

The answer to the question “How to Cause the Change?” is provided by the final implementation plan, where stakeholders help to sequence all the inputs from Steps 3 and 4 initially into a simple implementation roadmap showing the sequence and responsibilities for implementing the major injections. They then translate this high-level roadmap into a detailed S&T tree that will be used for validating the analysis and solutions (to close the gap and keep it closed) with other stakeholders and other levels as well as for planning and monitoring the execution.

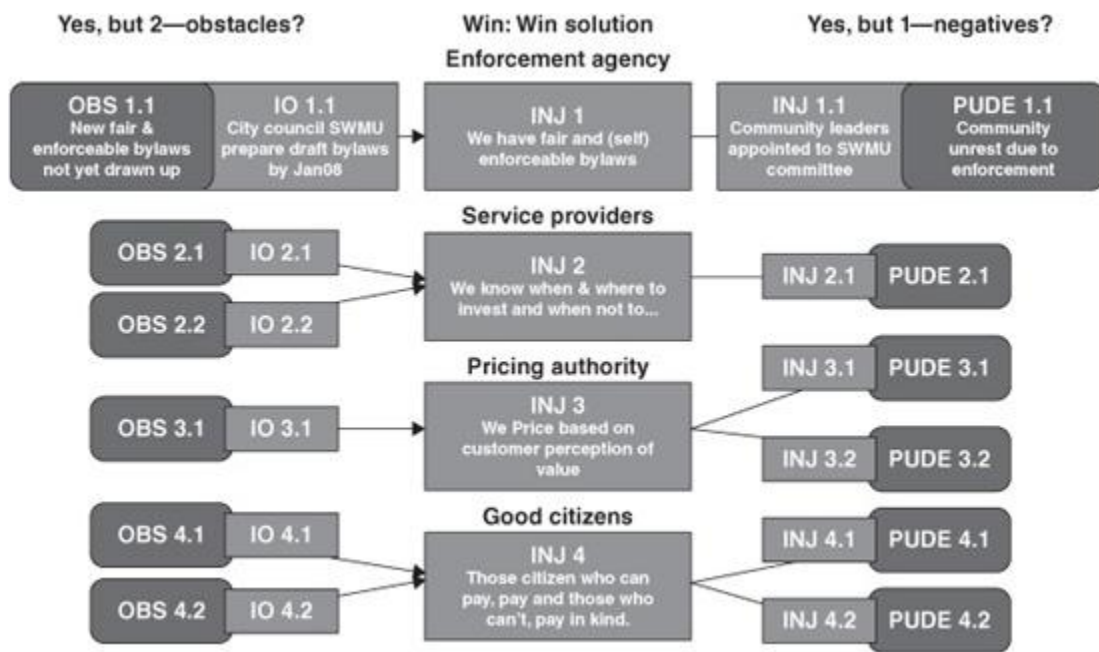


FIGURE 16-12 New simplified framework to enable participants to identify and overcome “Yes, but...” reservations.

Step 5—Getting Agreement on How to Measure the Change and Achieve a POOGI?

The last day was scheduled for agreeing on exactly what changes each stakeholder will make or contribute (especially those that must be implemented in the next three to six months), how the impact of these changes or contributions can be measured, and what additional capacity building and/or IT system support each stakeholder will require to achieve a POOGI within their part of the system or organization. The TOC expert also reviews with the groups how the equivalent of T, I, and OE can be used to judge the global/system impact of all decisions through a number of examples to identify past changes that should have been approved that were not or that were approved and should not have been.

There are two types of measurements needed for monitoring execution. A primary or “system status” measurement is used to measure whether the gap in the system goal is reducing, remaining the same, or increasing (i.e., gap between waste created and collected). The second type of measurement (secondary) would monitor whether the agreed injections and detailed requirements were being implemented and whether they are sufficient to achieve the primary measurement. These measurements would then be reported at a weekly operations meeting and monthly steering committee meeting with representatives of all stakeholders that would need to collaborate to identify recovery plans in the case where gaps are not closing at the desired rate.

Getting Stakeholder Feedback from a Strategy Session

On completion of the formal analysis steps, the most senior person of the organization, with assistance from the facilitator asks each participant to share their “what went well,” “what can be improved,” and “how I can apply the lessons learned in my day-to-day job and even in my life.”

We normally recommend two feedback mechanisms to understand how participants really experienced the workshop—one done on completion of the workshop in the open forum and another done a week after the session where respondents can choose to give feedback anonymously.

The responses of participants received from all four cities were very positive and showed that many of the paradigms and TOC processes did “stick”—even after only five days. The other positive feedback was from InWEnt, who had commissioned the TOC experts and other donor organization stakeholders.

A sample of open comments from participants captured by the InWEnt scribe after the first two 5-day Constraint Analysis Workshops in cities A and B:

- “I’m very grateful for my being able to fully participate. We have seen that with TOC, Goals are possible (even if they look impossible) and that the key is to find win-win solutions and focus on the weakest link”—**City Councilor**
- “I did not expect much but now have seen it is possible to work with all the stakeholders as our real goal (to improve the lives of everyone in our community) is the same”—**Private Contractor**
- “Discovered that we had a very big gap in our system. Previously thought we can only close it by getting more/new equipment. Now learned that I get more with the same equipment using the TOC approach”—**Private Contractor**
- “I learned that the ‘cost of doing nothing’ could be very high. In the health sector the ‘cost of doing nothing’ is a matter of life and death. TOC showed that we can do more with same resources just by better exploiting what we have”—**Doctor**

- “I learned the need for planning and that the impossible can be possible”—**Chairman of a Community Settlement**
- “TOC can even help to address the political issues”—**Councilor**
- “Not wait for more/new resources but start with what we have”—**Habitat Officer**
- “My expectation was low at the start of the workshop, but have learned the need and benefit for everybody’s participation using this TOC process”—**Council officer**
- “I learned that TOC is a window that gives a chance to look at the challenges we face in a different perspective and learned the importance of defining the problem precisely before dealing with it”—**Deputy Mayor**
- “Learned that the best solution comes by asking the right question”—**Councilor**
- “We like TOC because it gives us the skills to achieve ‘Continuous Improvement.’ It also provides a common language and process that will ensure we can achieve and maintain peace amongst the stakeholders. If everyone is exposed to TOC, they will know how to deal with Conflicts, Constraints and Yes, buts etc.”—**Mayor**

Probably the most telling was received from one of the mayors with which he closed the strategy session for his city:

“Today, I am so proud to be an African. What this team has achieved in only a few days and the spirit in which it was achieved was truly inspiring and should serve as a benchmark to the rest of the Public Sector of what is possible if the right people are put in a room together, reminded that we have a common goal and that it is always possible to find win-win ways to break

conflicts and overcome constraints, especially when lives depend on it.”

[Figure 16-13](#) shows photos of the active participation obtained in these workshops.



FIGURE 16-13 Selected photographs from the pilot workshops.

After the success of the first application, the concept of the 5-Day Constraint Analysis Workshop was used between April and August of 2007 in three additional cities. [Table 16-3](#) below shows the results that were achieved in each of the four cities.

Current Status of Pilot Projects (by the End of 2009)

The primary objective of the 5-Day Constraint Analysis Workshops was to test whether it was possible for each of the cities to achieve, in 5 days or less, the active contribution and consensus on each of the 5 change questions. The first application validated this hypothesis

and, as a result, the same methodological approach was followed in the 3 additional cities.

Although the activities at one of the cities did not continue due to another project that absorbed the manpower of the local Waste Management Unit, the 5-Day Constraint Analysis Workshops have led to a new quality in consensus building among key stakeholders and strategic planning.

After having led a two-year process of building up local capacities for improving municipal solid waste management by applying TOC combined with specialized waste-related technical knowledge, Michael Funcke-Bartz from InWEnt reports the following achievements of the TOC-based working process (scale differing from city to city):

- Waste management steering committees consisting of all stakeholder representatives have been established to manage the implementation process and to develop additional TOC solutions for obstacles identified during implementation.
- Official support from the political decision-makers was achieved that is a prerequisite for sustaining such a TOC process and implementation.
- Local junior TOC experts have been trained that are now able to facilitate meetings of Waste Management Steering Committees and with stakeholders applying key elements like obstacle based planning, conflict resolution, Throughput Accounting.

No.	Success Criteria	City A (Baseline)	City B	City C	City D
1	Agreement on the goal and gap	Yes	Yes	Yes	Yes
2	Agreement on the problem, system constraint, major UDEs, and core conflicts by stakeholder	Yes	Yes	Yes ^a	Yes
3	Agreement on the direction of solution to better exploit/elevate constraint—the injections needed to break core conflicts for each stakeholder	Yes	Yes	Yes	Yes
4	Agreement on the most significant negative side effects and how to prevent them	Yes	Yes	Yes	Yes
5	Agreement on the implementation obstacles and how to overcome them	Yes	Yes	Yes	Yes
6	Commitment by top leadership to support fully TOC initiative and take ownership to drive implementation	No ^b	Yes	Only for pilot sites	Yes

^aThe pilot projects in this city included representation from national politicians and therefore the UDEs and core conflicts of politicians expanded the previous four core conflicts found in other cities to five.

^bAlthough there was much excitement from all stakeholders (including the leadership of the city council) after the workshop, only some stakeholders implemented their injections because changes to the top leadership resulted in withdrawal of their support (and funding) for this initiative.

Source: Post-workshop debriefing results (by survey completed by all stakeholder representatives).

TABLE 16-3 Comparative Results of Outcomes from the TOC Workshops for the First Four Cities

- The joint working process has improved the understanding and cooperation between local government, private contractors, and informal waste collectors and has led to a paradigm shift that recognizes that important improvements can be achieved by better exploiting the available physical, financial, and human resources.
- In one case, the local government's financial conditions were so difficult that it was not even able to cover the running expenses for salaries. In this case, the TOC analysis helped to develop measures to increase revenues to overcome this constraint.

- Pilot projects were developed for servicing low-income areas through informal waste collectors.
- The legislative environment was analyzed and, in addition to necessary changes to laws and bylaws, the realization grew that existing legislation very often could be used by arrangement with the courts to resolve issues, previously deemed impossible.
- There has been evidence that waste collection has been increased due to better “Constraint Exploitation Rules.” For example, the number of trips done by Waste Collection Trucks increased by between 30 and 100 percent. Also, private contractors have reported significant increases in their revenues (due to more trips and higher levels of payment by the communities they service) and the community representatives have reported a visible decline in waste previously dumped, burned, or simply not collected.
- Examples were found where the drive to privatization resulted in a dysfunctional system. The elements of effective and efficient privatization efforts were examined and a functional systems model developed which could provide orientation to local governments on separation of functions while maintaining overall responsibility for waste collection services.

The remaining challenges include:

- Despite significant revenue increases to both the City Council and private contractors (typically up by 25 to 50 percent), in some areas there is still not enough revenue (not enough citizens and businesses paying) for the system to be financially sustainable. The injection for improving viability is that private contractors operating in these areas will be assigned additional areas (outside

their normal licensing area) to make them viable. This change of the City Council taking responsibility for ensuring the viability of the private contractors has made a significantly positive impact on re-establishing the trust and spirit of win-win (and recognition that a lose for one is a lose for all).

- In some areas, contractors were complaining that even though the fee has been reduced some citizens still do not pay because enforcement is not yet in line with expectations. A potential injection being tested is to put water, electricity, and waste removal fees on one bill. If citizens do not pay, you cut off all the services. In the past, this was not possible as privatization commonly apportions water, electricity, and waste services to separate companies who had no interest in combining bills. The City Council and private contractors believe this injection can potentially have the biggest impact on long-term system viability.
- There has been an unwillingness to put measurements and short-term targets in place. There is frequently a fear within the City Council to commit to achievement of specific targets (for political reasons). The injection here was to show how critical feedback measurements ensure that scarce resources are not wasted on areas doing well or that it should be moved to areas where the gaps are growing. The primary measurement that will be introduced in 2009 will be “Tons of Waste Collected.” For this, infrastructure such as weighbridges, etc., would have to be upgraded or replaced. Funding for this, based on the likely benefit of doing this, has already been secured in a few of the cities.

Future Application of TOC within the Public Sector

The projects carried out by InWEnt have not been the only attempt to incorporate TOC's holistic approach as part of Capacity Building initiatives to secure and sustain improvements within the public sector. As a spinoff effect of the initial projects, a national Ministry of Agriculture requested assistance from two of the TOC Expert facilitators, Professor Antoine van Gelder and Barry Urban, to use TOC to assist with prioritizing and synchronizing initiatives needed to close the gap between demand and supply of basic food groups.

Another example comes from the UN Development Program's Southern African Capacity Initiative. Governments have been struggling to meet agreed service delivery targets. After receiving advanced training in the TOC concepts, analysis processes, and logistical solutions the team has carried out their own pilot studies which resulted in them incorporating key TOC concepts and processes within their engagement process. Excellent work is also being done to expand the use of TOC within the health care and education systems around the world (www.toc-healthcare.com and www.tocforeducation.com).

Specific Lessons Learned from All the Public Sector Pilots

The TOC Analysis Process Does Work within the Public Sector

During the 5-Day Constraint Analysis Workshops in the InWEnt project, the TOC expert team showed that with a simplified TOC-based conceptual model—the Constraint Analysis and Solution Development Process—we could, despite the added complexities and uncertainties of the public sector, achieve the objective of getting all stakeholders to agree, in five days or less, on the answers to

the fundamental questions of Why Change, What to Change, to What to Change, How to Cause the Change, and How to Measure and Continuously Improve.

There Really Are Similar Types of Core Conflicts and Constraints for the Same Type of System

The expectation after the first workshop with City A was that we would find very similar, if not identical, core conflicts and constraints elsewhere because the same (erroneous) rules were used to manage these systems. This expectation was validated as the process was rolled out to the other three cities. The only observed deviation is when additional stakeholders (e.g., politicians) are introduced. This adds one more core conflict (for that stakeholder) that needs to be broken. We will soon get to a point where “analysis/discovery” becomes more of a “validation” against known causes and solutions. Already, after the second workshop (when we proved that the same generic core conflicts existed), I could hand over the facilitation to a cofacilitator, something that would not be possible so quickly if a “new discovery” was needed each time.

It Is Possible to Train-the-Trainer with the New Simplified TOC Process

One of the major concerns after the first successful five-day workshop in City A was whether someone else could take over the facilitation, especially someone without years of TOC experience. The realization, therefore, that we really did identify generic core conflicts at City A was very good news, since the scalability of this type of initiative is to a large extent governed by the availability of experienced TOC experts (that can identify new patterns) that are also good facilitators—a resource that is generally quite scarce. By showing that the major emphasis will move to validation rather than analysis and discovery, it meant that good

facilitators could be trained relatively quickly through “train-the-trainer” programs especially if we can provide such facilitators with detailed Strategy & Tactic trees. Nevertheless, working in the multi-stakeholder environment of the public sector requires facilitators who are sensitive and understand and accept the specific dynamics and decision-making structures.

Successful Implementation Requires Follow-Up and Follow-Through

In order to be sustainable, Capacity Building on the municipal level requires (like any important change initiative) active follow-up (prioritization and monitoring progress) and follow-through (identifying and overcoming implementation obstacles) by both stakeholder representatives and for example the TOC experts. This process should go together with enhancing the national financial institutions’ capacities to identify and prioritize investments in basic infrastructure that also follow constraint oriented thinking. The same is true for the national development plans. We also learned that the change management process needed to transform a dysfunctional system into a functional system can take longer than the normal term of the local governments’ leaders. This implies risks of discontinuity that can even stop the process. The multi-stakeholder steering committees that have been established in the participating cities have been trained in TOC thinking and are supported by local TOC junior experts capacitated during the project. They are an important element that contributes to securing long-term sustainability of the process. These Steering Committees are crucial and should be constituted by “self-selection” by all participating stakeholders at the end of the 5-Day Constraint Analysis Workshop. This committee should be responsible for prioritization and monitoring progress (i.e.

serve as the Program Audit Committee), institutionalizing the new changes and help contributing stakeholders identify and overcome implementation obstacles. Through these projects, it has also been shown that S&T can be used as an effective internal and external auditing tool and for testing key assumptions around necessity, sufficiency, and sequence.

The Importance of Finding and Sharing “Pockets of Excellence”

The pockets of excellence that do exist within the solid waste management units and private contractors and community-based enterprise must be identified and leveraged. This is important both to provide benchmarks that can be turned into best practices for rollout as well as to provide a way to quantify the inherent potential that can be unlocked if we can move from competing for survival to collaborating for success.

The Key Is Finding Out How to Change and Then Sustain The Desired Behavior

In May 2005, an article appeared in Fast Company titled “Change or Die” where the author, Alan Deutschman said that, “All leadership comes down to this: changing the behavior of people.” He claimed later in a book by the same title (Deutschman, 2007) that in the United States, 80 percent of the healthcare budget was consumed by five behavioral issues (smoking, drinking, eating, stress, and too little exercise) by a relatively small percentage of the population. This has been known and validated by science for many years. Still, the way to change these fundamentally damaging behaviors are still escaping policy-makers and doctors, even to the extent that most people when faced with the decision to change (their lifestyles and

behaviors) or die, they end up making the wrong decision by not changing.

In managing organizations, we know that people at all levels find it hard to change. Their awareness of the *negative consequences of not changing* and/or the *benefits of changing* is seldom sufficient to trigger the necessary change. This is mainly due to an awareness and often, an over exaggeration of the *potential negatives of changing* or *benefits of not changing*. For successful changes to occur, we need to affect the psychological, emotional, and spiritual dimensions, not just thought. Helping people to reframe the concepts that influence their thinking is the key. This is done by creating a new narrative that's simple, easy to identify with, and emotionally resonant.

TOC, if used properly, can play such a role as it can help stakeholders consider all four aspects (the positives and negatives of *changing* as well as the positives and negatives of *not changing*) related to a change in a holistic way. The analogy of focusing all stakeholders to contribute to strengthening the weakest link is a simple, easy-to-identify-with concept, and if combined with the commitment from all stakeholders that win-win-win is the only sustainable solution, can provide the emotional resonance to inspire contribution.

Where Are There Still Gaps or Complexity in the TOC Roadmap and Process?

Feedback from the InWEnt and UNDP projects reminded the team that there are two important steps in the analysis, which should not be skipped:

1. Recognize past achievements (but then show the large current and future gaps remaining and the consequences on all stakeholders of not closing it).

2. Validate the impact of current/planned strategies (that is, can they help break core conflicts/constraints and how will T, I, and OE be impacted) or will they be in conflict with the new strategy? To release the necessary capacity and/or prevent conflicting priorities, these existing programs should be listed and reviewed on the last day since frequently they fall into the category of “what should be STOPPED so we can START doing what is really needed now.”

If either (1) or (2) are not done, the workshop will not achieve its objective of reaching full consensus and commitment from all stakeholders on Why Change, What to Change, What to Change to, and How to Cause and Measure the Change to Close the Gap.

Future Research

As mentioned before, managers within public sector organizations, like their counterparts in the private sector, know that the fact that we reach full consensus on what needs to change and how and by whom the change will be implemented does not mean it will be done. To ensure proper execution, we need a mechanism to enable managers at all levels to follow-up and follow-through. Edison’s famous quote, “Vision without execution is a hallucination,” is a reminder of this simple fact of life.

One way to ensure that good ideas are turned into results is to capture the outcome of a TOC analysis workshop in the powerful S&T format, described in [Chapter 15](#), and then to go one step further by using the S&T tree to develop the implementation plan and manage and monitor execution. [Figure 16-14](#) shows one of the S&Ts created in Harmony ([www. goldrattresearchlabs.com](http://www.goldrattresearchlabs.com)) by the InWEnt/TOC and Waste Management expert team for a Solid Waste Management project.

Holistic TOC Implementation in the Private Sector

Where the public sector is typified by bureaucracy, the private sector is typified by competition and immediacy. Companies prosper from their ability to innovate and renew. As soon as a for-profit organization achieves significant levels of profitability, competition will enter the same market sector. All for-profit strategies contain two simple requirements:

1. Defend what you have for as long as possible. Traditionally, this is achieved through intellectual property rights, embedded technology, skills that are difficult to replicate, and price manipulation.
2. Change and improve more rapidly than competitors. Again, this is achieved through successful technology, R&D, mergers and acquisitions, and product or service enhancements.

Before the advent of TOC, the ability to manage change itself as an explicit and discrete strategy had rarely been considered to be the key to rapid and sustained business growth.

It is frequently easier to focus on the first option highlighted above because most of the variables are known, costs can be calculated, and outcomes are fairly predictable. The second option is more difficult to execute and inherently more risky. However, only the second option ensures long-term prosperity. Therefore, the *rate* of improvement is the crucial variable in long-term success.

We advocate a new frame of mind: one where the *process of change* in itself becomes the central and accepted function of management. This is not about incremental change, but about large, system-level step change. Top managers who understand this have the greatest potential

of success. This type of change is not easy to do because it involves significant work and risk, but has the potential to deliver impressive results as described in the case of First Solar Inc.

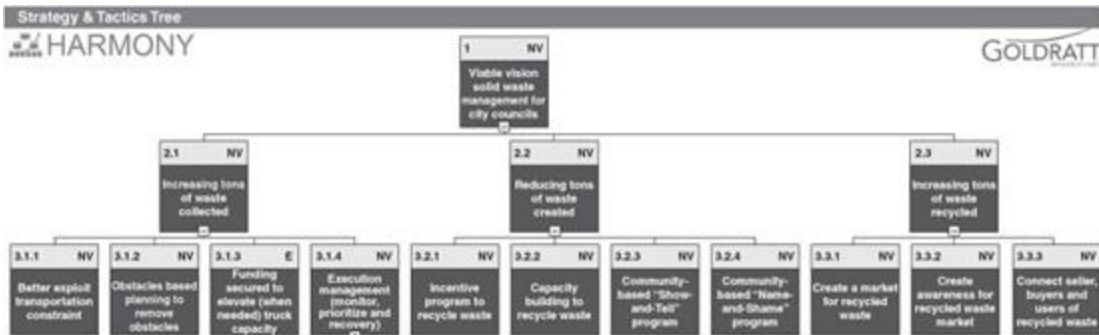


FIGURE 16-14 Example of a public sector SWM S&T created in harmony.

The Birth of First Solar Inc.

First Solar’s origins can be traced back to the oil crisis in the 1970s. At that time, a well-known American entrepreneur, the late Harold McMaster, realized the need for low-cost renewable energy to mitigate the risk of fossil fuel supplies. With his experience and financial success in glass production and related manufacturing equipment, Harold rounded up 57 investors to fund Glasstech Solar in 1984. He wanted the investors to help fund his vision of producing solar panels so efficient that 2000 square miles of solar panels located in the Arizona desert would satisfy the entire United States’ need for heat and light.

In 1987, with the sale of Glasstech holding company, Harold kept ownership of the solar business. With 47 of the original investors and \$11.5 million, he continued to pursue his vision. Based in Toledo, Ohio, the company focused on a unique photovoltaic semiconductor process—that of

depositing thin-film semiconductor on a glass substrate. The initial focus was on amorphous silicon as a semiconductor but, by 1990, after burning through \$12 million, the company was running out of money.

Under pressure from Jim Nolan, Vice President of Operations, Harold decided to change course. Glasstech Solar became Solar Cells Inc. (SCI) and Harold offered to pay back any of the investors who wanted out. He raised another \$12 million and started again with a relatively unknown and risky photovoltaic semiconductor material—cadmium telluride (CdTe). Cadmium telluride had an ideal band gap, a simple structure, required no doping, produced films that bonded well, was tolerant to impurities, and could be constructed from mostly readily available materials. The crucial missing piece was the ability to manufacture CdTe solar panels in large quantities at a low cost. Given that the entire global demand for solar panels was only 125 megawatt (MW) at that time, the strategy to develop large-scale manufacturing capability was a huge gamble in the face of proven and established crystalline silicon technologies. To put that in perspective, the 2008 demand for solar was 5950 MW. Racing against time and very significant cash constraints, Jim Nolan and his team invented a unique and innovative Vapor Transport Deposition technology that provided the breakthrough to large-scale manufacturing. By 1993, the company had built a pilot production machine, driven by a flurry of technology breakthroughs. Cash was so tight that employees bought discarded school furniture at auction to use as office furniture.

In contrast to Harold's 500-MW dream plant, Jim wanted to commercialize the technology as soon as possible. Trying to find buyers for a completely new technology in a market that still considered photovoltaic energy a curiosity proved a real challenge. By 1993, the plant was producing 0.2 MW per year—a far cry from Harold's 500-MW vision. By 1995,

five of the seven key managerial and technical staff had resigned, considering SCI too risky to sustain a career. Harold sold yet more shares, forced to buy most of them himself because the high level of perceived risk deterred even the most daring investors.

By 1998, SCI had burned through \$35 million without turning a profit. Time was running out.

Despite living on the ragged edge of failure for so many years, the dream came to life again when the company reached a new milestone—the production of a 60.3-W module with 8.4 percent efficiency. They knew there was gold at the end of the rainbow.

By 1999, the lone manufacturing line was producing 20 panels a day, still fraught with quality problems. Adding to the systemic risk was the demise of BP Solar, the only other major company that had been working on commercializing CdTe. It dealt a huge blow to investor confidence and left SCI as the lone entity in the United States and a dark future.

Enter True North Partners, a private equity firm backed by the late John Walton, son of Sam Walton, the creator of Walmart. Mike Ahearn, a partner at True North, saw the long-term business potential in thin-film solar technology and convinced John Walton to invest in SCI, providing some \$10 million in operating capital.

The investment opportunity was made interesting by the following:

- Semiconductors were manufactured without the use of clean rooms.
- The manufacturing process was self-contained and end-to-end.
- Widely available materials were used other than for the semiconductor itself.
- The manufacturing process was low cost and easily scalable.

With the new investment in the now-named First Solar came Mike Ahearn's hard-nosed business skills and relentless drive for bottom-line results, which unfortunately did not fit well with the pioneering culture of "turning knobs and seeing what happens."

Management found it difficult to adapt to True North's expectations and, in 2003, True North Partners bought out the entire shareholding, gaining full control of First Solar's destiny. From that moment on, the company blasted off on its meteoric rise, culminating in its inclusion as an S&P 500 company in 2009.

One of the first actions Mike Ahearn took was to assemble and integrate a superb top management team. This entailed the painful hiring and firing of managers until the best team was in place. To bring discipline to experimentation with manufacturing processes, Taguchi's Quality Engineering approach was introduced. To reduce process variation, Six Sigma was introduced, and to maximize Throughput, TOC was introduced. This was executed under the leadership of the new COO, George "Chip" Hambro.

As the manufacturing process matured, the constraint moved to the marketplace. In November 2002, Kenneth "Ken" Schultz joined the company as Vice President of Sales and Marketing. Having worked with John Walton in another company in the past, Ken was faced with the task of finding a market for the avalanche of solar panels that now started pouring off the much-improved manufacturing line.

The company had been searching for new applications for solar panels, including solar-powered refrigeration and similarly powered water pumping installations, but none had demand that would match First Solar's growth objectives. As luck would have it, two events on the other side of the world were about to change First Solar's destiny forever.

The German federal government introduced legislation to provide financial incentive for the use of renewable energy. This triggered extraordinary interest and growth in the use of renewable energy, including photovoltaics (PV), just as First Solar's low cost, high volume manufacturing process was coming to life.

The second event was the demise of Antec Solar after multiple bankruptcies. That company had been the only other solar panel manufacturer using CdTe as its semiconductor material. That left First Solar as the only company able to manufacture low-cost solar modules in large volumes. Trade press banner headlines described First Solar as "The Last Man Standing," predicting that the CdTe technology was all but dead.

This helped detract potential competitors from developing the same technology, opening up a precious window of opportunity where First Solar could perfect its low-cost, highly automated manufacturing technology.

With the advent of the renewable energy feed-in tariff, a number of project development companies also emerged in Germany. They found the large-scale project development opportunities, and coordinated the finance, site permitting, and construction. Their entrepreneurial, high-risk culture meshed well with that of First Solar, and First Solar forged close ties with key players to secure a strong foothold in the nascent market.

Throughout the period of torrid growth, Mike Ahearn steered the corporate strategy by three simple rules:

1. Set goals that were clear, measurable, and challenging.
2. Systematically identify and remove constraints that prevent the company from achieving those goals.
3. Drive execution excellence.

This simple, but by no means easy approach had a profound impact. It resulted in the exponential growth of annual manufacturing capacity from 25 MW in 2005 to more than 1000 MW in 2009, a 40-fold increase in four years.

The increase in capacity was matched by a commensurate gain in output as shown in [Fig. 16-15](#). The leftmost vertical axis represents annual module production, while the rightmost vertical axis represents module efficiency. The horizontal axis is the period from 2001 to 2007.

Mike Ahearn also re-engaged Jim Nolan, the stalwart VP of Operations from the early days, appointing him to the First Solar Board of Directors to secure access to his deep knowledge and understanding of the company. Throughout this time, John Walton remained a committed provider of capital, investing some \$165 million up to the IPO in November 2005. That investment turned into a public company with a market capitalization of some \$12 billion.

By 2009, sales had grown to \$1.2 billion and net profit had grown nearly 10-fold. The First Solar management team gained the respect of Wall Street, with its enviable record of delivering on its promises. One analyst wrote: "I have gained much more appreciation for the caliber of First Solar's management team. This company has executed at an extremely high level and is well prepared to face future challenges."

One Merrill Lynch investment analyst figured out what was really contributing to this enviable track record. In an investor note dated June 2009, he wrote:

Corporate philosophy revealed

Mixed in with the roadmaps on cost/watt and conversion efficiency at the analyst meeting were telling indications of First Solar's business approach. Although financial performance eventually dictates stock price, our view is that financials are a lagged reflection of philosophy, strategy, and execution. Management's goal of building a company to

last 50 to 100 years may seem like hubris but is the right approach for an early leader in a potentially huge market.

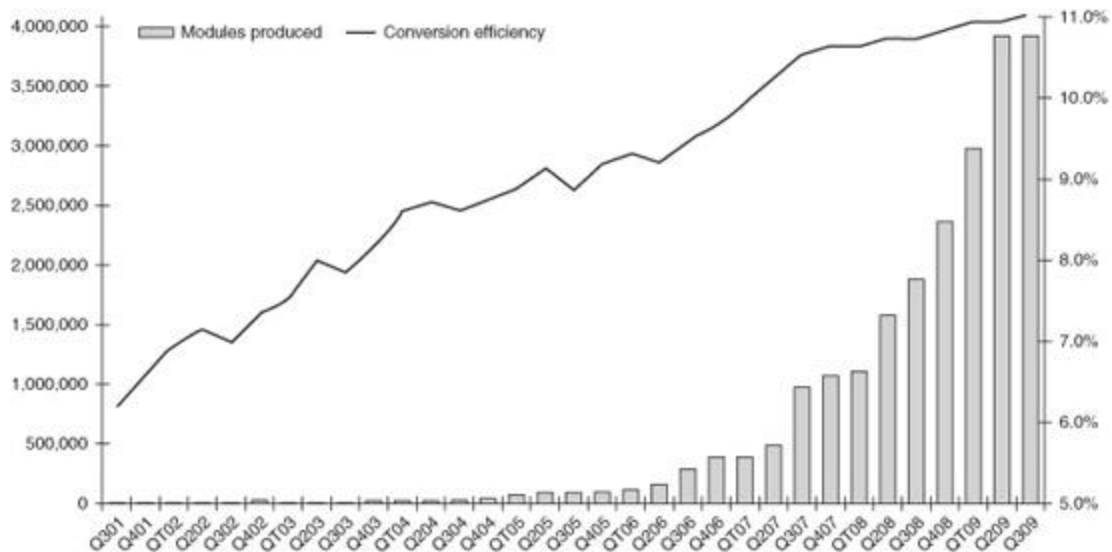


FIGURE 16-15 Exponential growth of First Solar’s annual Throughput.

Theory of constraints explains corporate actions

Management alluded a few times to using a “Theory of Constraints” in running the business. Theory of Constraints is a management approach introduced by Dr. Eliyahu Goldratt in 1984 that says organizations can be measured by Throughput, Operating Expense, and Inventory. The company’s acquisitions and other actions have been aimed at removing constraints throughout the solar value chain.

What follows is a description of this remarkable holistic application and implementation of TOC.

Theory of Constraints Contribution to First Solar’s Success

At the outset, TOC is one of a set of tools and approaches introduced to drive performance improvement in First Solar. The impact of TOC (and similar approaches) is determined

entirely by top management's attitude to change and its assumptions about limits to sustainable growth. It is only an enabler of extremely capable managerial skill and orientation. Although TOC started out as a set of Throughput-focused applications, TOC grew to become a business management philosophy embedded in decision-making at many levels in the company. Therefore, the key question that this case study serves to answer is "How does management turn a set of Throughput-enhancing tools into an embedded Throughput-driven culture?"

To appreciate the significance of TOC as a business management philosophy, we should place it in context of the 21st century global economic and business environment.

The advent of the Internet and the massive increase in the availability of information has made the world a much smaller and connected environment within which to conduct business. Any comparative advantage that a company may have has a much shorter and precarious life. In order to stay ahead, management is more pressed to drive continuous improvement faster and more predictably.

There is a very significant risk of applying resources to low leverage improvement opportunities, which may slow growth. You may recall Microsoft's spectacular failure to grasp the impact of the Internet when Netscape became the dominant browser, only to have the error repeated when Google came to dominate the search engine space.

By truly understanding the constraints that limit growth and profitability, scarce resources can be focused on the most valuable points of leverage. First Solar top management deliberately and continually assess the internal and external business environments to identify the constraints to rapid growth and to focus the organization on leveraging those constraints.

Organizations that consider the inward focused, internal efficiency of individual resources to be the primary cause for

success are frequently blindsided by changes in behavior of the large systems within which they exist.

Conversely, those that consciously, and with discipline, take advantage of such system changes are bound to be more successful in the long-term. TOC accomplishes exactly that. First Solar's stellar success can be attributed also to the diligent analysis of internal and external system-level constraints.

Building the Foundation

Three aspects are crucial to elevate TOC from a defined toolset to a deeply embedded culture: *Throughput World versus Cost World, the importance of systems and constraints, and the importance of a dedicated and motivated TOC subject matter expert.*

Throughput World versus Cost World

This very subtle, but crucial aspect affects the entire organization.

Cost World companies are characterized by an inherently defensive mindset. There is a general feeling that the company always has its back to the wall. Expenditures are subject to prolonged scrutiny, employees are constantly in a state of disquiet, many functional silos exist, and frequently there is low-level internal warfare between management and employees. In biblical terms, there is a general feeling that it is better to hide the talents than use them.

In contrast, Throughput World companies are characterized by a belief that success is only limited by management's ability to identify and remove system-level constraints. Goals are merely milestones on a continuing journey to higher levels of success and not the annual justification of contribution. Teamwork is understood to be

essential, embraced to help remove system constraints, and used to drive large, cross-functional improvements. Where traditional improvement approaches are focused on optimizing the existing system components, TOC assumes that large system change is essential for rapid and sustained growth. For this to happen, a consistent value system and decision framework must exist across the organization.

Over a period of six years, nearly all business decision-making employees in First Solar were educated in the fundamentals of TOC. Using the Socratic teaching method and simulation models, participants were taught the principles of a Throughput-driven organization and a simple decision-making model applicable across all functions and systems. The core concepts are:

- *Maximize Throughput.* For every business system within the organization, the requirement is straightforward. Continuously improve the Throughput rate and value that systems deliver to First Solar. Conceptually, Throughput has no limit.
- *Minimize Inventory.* Lower inventory levels enable higher Throughput. Inventory should be reduced to a level just above where Throughput is compromised.
- *Minimize Operating Expense.* Operating Expense reductions are judged by the impact on Throughput and should be reduced to the level where Throughput is not compromised.

Systems and Constraints

Throughput World companies succeed in the long-term by relentlessly driving improvements through all of the

business systems. This can be accomplished by:

- Identifying, describing, and analyzing all the core business systems in the organization.
- Identifying the constraints within those systems.
- Removing the constraints, making sure that no unintended consequences prevail (through better exploitation or elevation).
- Reconfiguring functions, jobs, responsibilities, resources, and measurements to sustain a higher level of performance (subordination).
- Repeating the process.

Employees who have worked in Cost World companies find this to be both energizing and exhilarating, but also more stressful because significant change is the order of the day.

The last point serves to highlight the critical importance of top management's deep understanding and commitment to being a Throughput-driven company.

In this context, the successful and sustained use of TOC in First Solar during the past six years is due entirely to the support and commitment of the CEO, Mike Ahearn, and the President, Bruce Sohn. They share a deep understanding of the subject and have firsthand experience of its impact on success.

Deployment of Dedicated TOC Subject Matter Expert

The third important aspect to embed TOC deeply in the culture of the organization is the deployment of a motivated and dedicated TOC subject matter expert or team. Ray Immelman, a TOC pioneer and now Vice President of Strategic Services at First Solar, leads such a team. That

team acts in a staff role to top management and shares one extremely important character trait essential for success. Despite the stress of simultaneously changing politics, people, and systems, they never let go or let up. Their motto: Don't blink!

A strong TOC culture offers more value than just improving the internal functioning of the organization. When First Solar initiated its IPO discussions with Goldman Sachs in 2005, the Throughput-driven culture was presented as an important element in the company's success. In addition, when capital equipment vendors became a constraint to First Solar's growth, the company exerted pressure on them to become more Throughput-driven. Subsequently, those vendors exhibited similarly impressive growth trajectories and in turn exerted pressure on their vendors to do the same.

Unbolting the Existing Systems and Measures

Since the Cost World mentality is driven by traditional assumptions and operating paradigms, companies implementing TOC should not underestimate the effort required to identify the measures, information systems, policies, and business rules that are in conflict with the requirements for establishing a Throughput-driven company.

Luckily, in the case of First Solar, there was very little to undo since Solar Cells Inc. was eager for a way to shake off its hand-to-mouth existence. Still, it took a number of years before the culture truly changed. The rallying cry that sustained the implementation effort was "Throughput." That single word is the fulcrum around which the organization's culture was built.

It also provided the measure against which all existing systems are evaluated and changed. If systems, policies, and rules are Cost World oriented, they are treated with due

care to ensure that they don't become constraints in and of themselves. For example, the CFO once wished to introduce a new set of accounting rules. The CEO flatly refused to implement the recommendation, insisting that it would introduce a policy constraint, which would affect the Throughput of the organization. This has become a norm for many on the management team.

Building on Early Success

The old adage that “nothing succeeds like success” is exquisitely truer in an environment where large system change is undertaken. Although early success in the use of *Drum-Buffer-Rope* (DBR) principles immediately improved the Throughput of the manufacturing process, the real watershed event was the construction of the first plant expansion. Early on, Mike Ahearn discerned that the existing manufacturing facility would not be able to scale to support the required growth, and decided that a new facility must be constructed. It would not just be an extension of the existing facility, but a new, state-of-the-art plant. The key manufacturing technology was all that carried over. Everything else was new—new equipment, new vendors, new plant design, and new IT systems. To complicate matters, the expectation was that it would be accomplished in record time.

The first project plan developed by the project team showed substantial completion in mid-2005. After introducing *Critical Chain Project Management* (CCPM), the anticipated completion was October 2004—a 30 percent reduction in the timeline.

Despite the project team being understandably apprehensive that this was not doable, they did complete the plant in that time frame.

The impact on team psychology was immense.

In contrast to the previous culture of hand-to-mouth existence, there was a palpable sense of significant achievement, which in turn created a belief that significant change was indeed possible. That watershed achievement formed the basis for First Solar's "Copy Smart" philosophy, which enabled the company to rapidly and predictably replicate plants much faster than the industry norm.

The factors that enabled that seminal event were:

- Intense support from top management
- Project leadership from Alan Henderson, Director of Corporate Engineering
- Implementing CCPM with software, extensive training, and active support
- The use of a dedicated Critical Chain project manager
- Progress reporting to board-level using *Buffer Management* (BM) as the reporting mechanism
- The use of a consistent language and a consistent process across the entire organization
- The use of the CCPM performance and improvement measures to gain even greater effectiveness
- A team-level response to buffer incursions
- Understanding the different cultures in each CCPM application area—IT is quite different from R&D or construction.
- Understanding the business drivers for behavior in each group—what drives the current behavior and how does this need to change?
- Supporting the decisions people have to make—ensuring that the behavior and information generated by CCPM contributes to the business decisions they have to make
- Providing strong technical support and reassurance—Critical Chain software is the only software that directly

touches people's work and corporate interaction (even on a daily basis)

Today, CCPM is in widespread use throughout the company, continuing to contribute tremendous success. For example:

- In 2007, as a new public company, First Solar had to comply with Sarbanes-Oxley financial control requirements. Using CCPM, the company successfully adapted all financial controls in less than one year, earning recognition from the external auditors, Price Waterhouse Coopers, for the fastest error-free implementation the team had ever seen.
- In 2009, the MIS organization doubled the number of projects completed in one year, using CCPM to manage their large project portfolio.

Implementing the Proven TOC Toolset

Once the Throughput-driven culture is established, and early success is celebrated, the remainder of the TOC toolset has fertile ground for implementation. What follows is a description of some of the embedded applications.

Drum-Buffer-Rope

When the first new plant was designed, BM was embedded and automated as best as possible. Pacing resources were identified, constraint and feeding buffers were established, and robots were deployed to execute most of the BM tasks.

Combined with the use of Six Sigma applications to eliminate process variation, line utilization is consistently in the high 90-percent range. One aspect that really helped to validate line performance is the extensive use of simulation modeling. Since TOC is all about the behavior of large systems, a key requirement for successful change is use of good simulation models. It provides:

- Insight into how the system will behave before actual implementation.
- Reassurance to all involved that intended changes will work.
- The ability to test all options, thereby ensuring strong buy-in.

Buffer Management

To leverage the application of DBR, the manufacturing team under the leadership of the plant manager, Todd Spangler, created the *TOC Information System* (TOCIS). This system enables management to monitor the performance of any manufacturing line anywhere in the world, in real time. It displays Throughput rates, current constraints, buffer inventories, and yield performance. All data are generated directly from the manufacturing lines and represent harsh reality.

In addition, the information is projected onto the factory walls in a simple graphic format for all to see. This has galvanized the attention of operators and supervisors alike, since team performance is unvarnished public knowledge.

The Role of TOC's "Thinking Processes" at First Solar

The TOC thinking tools are used to understand the changes to First Solar's current reality, to identify the core conflicts and assumptions that need to be understood, and to build and validate the logic for desired future reality. The First Solar management team uses these tools both as a day-to-day management tool in many instances, as well as to develop large-scale business strategy. In 2009, the company deployed the corporate strategy development and review process by having each operational unit present the current reality, injections, and future reality to the top management team. This enabled:

- Focused discussion around the high-impact issues.
- Deep common understanding of the way respective business systems behave.
- Alignment and coordination of work across business units.

A large contingent of upper- and mid-level managers has been educated in the use of the TOC TP, as has the CEO and some members of the Board of Directors. In 2009, the use of S&T trees was initiated to crystallize, condense, and communicate key strategic initiatives.

What Has Made TOC Work at First Solar?

There are a number of reasons why the holistic implementation of TOC has been successful and sustainable:

- TOC is simply a means to an end and not an end itself. The greatest benefit lies in creating a Throughput-driven

culture, supported by tools that make this a living capability.

- TOC provided a central concept around which a strong Throughput-driven corporate culture could be built.
- TOC had strong buy-in and long-term support from both the CEO and the President.
- Extensive exposure to the concepts and contrast of Cost World and Throughput World—this is a strong “tribal” motivator¹⁰
- Use of the tools in all possible areas helped define and build processes and systems around the concepts.
- Over time, success helped to build confidence in management’s capability, which drove an increasing virtuous loop.
- Small, successful initial applications built trust and confidence.
- Critical mass and bench strength in the TOC implementation team perpetuates the Throughput culture.

Recommendations and Summary

The aim of this chapter was to share important insights and experiences gained from both an internal and external TOC expert’s perspective as well as a generic process for implementing and sustaining TOC in a holistic way within both the private and public sectors.

With every implementation of TOC, whether done in one part of an organization (locally) or holistically (synchronized implementation in all the parts), we have the opportunity to learn both what worked and what did not work. One has to be careful to learn from each experience, but even more careful that one learns the right lessons.

Recommended Good Practices for Implementing TOC Holistically

The recommendations that follow are lessons learned that we consider key success factors or simply generic good practices for top management, internal and external TOC experts and TOC champions derived from experience with what did and did not work in practice.

Use the TOC 5FS, TP Roadmap, and S&T Tree

The best¹¹ place to start with a holistic implementation of TOC is to gain top management understanding and long-term support. This is best achieved by education in the principles and application of the generic TOC concepts or “gestalt.” Without a deep and visceral understanding and acceptance of the value of a Throughput-driven win-win-win culture, all else is short-lived. Once achieved, the next step would be to gain consensus on how to do just that by getting all senior management (and other key stakeholders) to contribute to constructing a business/organization strategy using TOC’s 5FS (applied to the organization as a whole) and the TP Roadmap to better leverage system constraints.

The S&T tree format can then be used to develop the full analysis (the what, how, and why related to each proposed change as well as the level and sequence of implementation). Experience has shown that the S&T tree is an excellent communication tool to validate with and get consensus from stakeholders on the necessary and sufficient changes, contributions, and implementation sequence.

In cases where a generic S&T tree exists that closely fits the organization, the focus should be on validating the core assumptions and making the necessary modifications to match the specific situation at that organization. In either

case, the S&T tree can also be used to plan, execute, and monitor the execution.

Changing Behaviors/Paradigms with Stories and Games

Changing behaviors requires a change in basic assumptions or paradigms through which we view the world. We recommend the following approaches for stakeholders to challenge and change their assumptions and decision-making framework.

- Using the five (traditional) limiting versus five (TOC) enabling paradigms ([Table 16-1](#)) to show that it is the way we deal with constraints, complexity, conflicts, uncertainty, and bad behaviors/choices (rather than these challenges) that will determine our success (or failure).
- Using the Throughput versus Cost World analogy to provide a simple, yet powerful explanation of the change in focus. Instead of incrementally improving the system by reducing costs everywhere, focusing on increasing Throughput (flow) through identifying and removing system-level constraints is much more effective.
- Use the multitasking game (Projects), dice game (Operations), P&Q game (Finance), and other TOC/systems-approach-based simulation games. These provide an effective and efficient (and fun) method for showing the potential gain in performance if one were to stop using local optima rules and start using TOC/system constraint focused rules. It also shows how frequently it is that assumptions block this potential, rather than the lack of resources or starting conditions, which are usually blamed for poor performance.

Important Role of a TOC Champion Supported by TOC Experts

For a company to start a TOC implementation, especially a holistic implementation of TOC, there must be an internal TOC champion who is well respected within the organization and supported by top management, with a critical mass of internal TOC supporters and (internal/external) TOC application experts. Internal TOC champions are people who are willing to take personal risks to move their organization, have adequate knowledge of the essence of TOC (what Goldratt calls the “gestalt”), and the ability to convince and motivate others (and off-course someone that will practice what they preach).

They relentlessly drive the culture change and ensure implementation support. We have observed that many TOC champions have limited success when they lack sufficient internal support, access to subject matter experts, or lukewarm top management support.

Finding and Leveraging Pockets of Excellence

Pockets of excellence exist in virtually all difficult or complex environments—whether it be in the private sector organizations that face the challenge of achieving exponential growth or simply surviving or the public sector organizations that face massive budget and resource constraints.

Every institution has its unique set of irrational and difficult constraints (Collins, 2006, 31), yet some make a leap while others, facing the same challenges, do not. Excellence or greatness is largely a matter of choice (where to focus) and discipline (taking the right actions and stopping the wrong actions).

In many cases, pockets of excellence can act as catalysts to identify what changes in rules/conditions contributed to stellar performance and possibly making the exception the rule.

Importance of a Simple Mechanism for Making Holistic Decisions

As long as decisions are still made based on many wrong assumptions engrained within traditional cost accounting, management decisions will continue to be in conflict with what is needed to support a holistic implementation of TOC. Implementing TA and specifically the mechanism to judge all decisions based on their impact on T, I, and OE is a simple and effective way to achieve better, faster global decisions at all levels. TA can help prevent devastating mistakes in decisions related to judging the financial performance of the organization as a whole or business units/departments, judging make/buy decisions, judging product/market/customer/project contribution, judging investments, and making pricing or budget decisions.

Using and Encouraging the Use of TOC's Thinking Processes

Management and employees face many challenges every day that could cause them to lose focus; challenges such as finding win-win solutions to day-to-day or even chronic conflicts, having to deal with half-baked solutions, dealing with firefighting due to gaps between responsibility and authority, dealing with resistance to change, and how to motivate and inspire teams to achieve ambitious targets. The TOC TP such as the positive and negative branches, conflict cloud/dual conflict cloud, PRT, and S&T trees provide managers with practical thinking process tools to help deal with these challenges. Acceptance is assured when managers are seen to be using this capability.

Importance of Roadmaps, Follow-Up, and Follow-Through

Getting people to take action is not a triviality. Appreciation of a body of knowledge such as TOC will not necessarily lead to a decision and action. The overriding factor in most of our decisions is not the expected benefits but the risks involved.

To mitigate the risks it is vital to have or provide a well-crafted, clear, and credible road-map backed by sufficient internal and external support. A well-structured S&T with valid assumptions provides a clear vision and garners strong consensus to action.

Unless there is follow-up and follow-through by leaders, the required actions either will not be taken or at best will not be sustained. Lack of follow-up and follow-through is rapidly interpreted that something is no longer a priority.

Importance of Aligning IT Systems with TOC Rules

Do not confuse simple with easy. Even simple rules can be difficult to implement, especially within large and complex environments. IT systems such as *Enterprise Resources Planning* (ERP) play an important role in helping organizations implement, apply, and sustain new Throughput-driven rules. They can help alter planning, execution, feedback, and control by identifying leverage points for continuous improvement. One way to ensure full alignment between the organizational growth/business strategy and the changes needed within the ERP and other IT systems is to use S&T as a tool to define and obtain agreement on detailed IT system changes required to support a new growth strategy. An example of the benefits of this approach was shown in the “One Simple ABB” case study where S&T was used to define and obtain agreement on the detailed ERP system (SAP) changes to support and

implement a set of simple TOC planning, execution, and continuous improvement rules within a large number of factories (Barnard, Rajaniemi, and Nordstrom, 2009).

Continuous Improvement and Auditing

Frequently, the business environment changes or reality proves (despite the best possible logical analysis) that one or more of the assumptions of necessity or sufficiency or inherent potential were simply not true. This calls for constant re-evaluation of assumptions about inherent potential, the system constraint that limits us from unlocking this potential, and the necessary and sufficient conditions needed to support this decision. [Chapter 15](#) of this Handbook provides a detailed description of designing and implementing a TOC-based continuous improvement and auditing mechanism and culture needed to support a holistic TOC implementation and to prevent or at least minimize typical management errors of omission, commission, detection, and correction.

Become the Change You Want to See

Our final recommendation to those considering the implementation of TOC in their organizations, especially to those struggling to get support or consensus, is simply to follow the advice of Mahatma Gandhi—“Become the change you want to see in the world.”

Start using the TOC TP tools and logistical applications in your area, and soon it will spread to other areas, sometimes well beyond the original area of influence.

Summary

Achieving a holistic implementation of TOC should start with setting a very ambitious target for achieving both sustainable growth and stability, not based on extrapolation from the past, but on the inherent potential. If an ambitious target already exists, it should be validated to ensure that it really would be a win-win-win for all stakeholders.

Once an ambitious goal has been set, time should be invested with all the stakeholders to really understand the gap (between actual performance and the goal) and also to ensure that all stakeholders understand the consequences (on the system as a whole using predicted effect logic to determine the likely impact on key stakeholders and using T, I, and OE to determine the financial impact) of both closing and not closing the gaps.

Next, the important decision has to be made on where to focus scarce resources to close the gap (to sustainably improve the flow of goal units) and, if not more important, to decide where *not* to focus. A proven way to get these answers is to use TOC's 5FS and TP such as the S&T tree. All such change initiatives of what to start and stop should be planned and buffered as a project (using CCPM) and should be pipelined based on agreed strategic priority and released only when the capacity to execute is available.

However, knowing where to focus is not sufficient. Top management must ensure that simple holistic rules for planning and execution (according to TOC's known logistical solutions) are in place and that obstacles/excuses for not starting the desired actions (e.g., implementing TOC planning and execution rules) or not stopping the undesired actions (that cause waste, local optima disharmony within organizations) are systematically removed. In addition, they must ensure that everyone has the means (systems) and ability (know-what, know-how, and know-why) for their required contribution, but also recognize contribution and success (and when convinced the means and ability were provided, hold those still not contributing accountable).

A fast (and reliable) feedback mechanism is required to continuously reduce the time to detect and time to correct for erroneous assumptions and for changing environmental conditions, especially for changes in the system constraint.

A POOGI mechanism such as TOC's BM should be implemented in each of the main functions to guide where further process improvement or capacity elevation is required most to maintain the balance between supply and demand.

Finally, don't blink. Persevere in the quest to achieve growth and stability despite all the corporate and individual idiosyncrasies.

References

Ackoff, R. L. 2006. "Why few organizations adopt systems thinking," *Systems Research and Behavioural Science*. 23:705-708.

Barnard, A. 2003. New developments and innovations in the Theory of Constraints Thinking processes, Paper presented at the TOCICO Conference in Cambridge, UK (October).

Barnard, A. 2008. Using a simplified Theory of Constraints approach to achieve more with less in less time within the public sector: A case study from Africa, Sun City, South Africa: SAPICS Conference Proceedings (June 29-July 2).

Barnard, A. 2009. How to identify and unlock inherent potential within organizations and individuals using a systems approach. PhD dissertation, Da Vinci Institute, Johannesburg, South Africa.

- Barnard A., Rajaniemi, K., and Nordstrom, F. 2009. Achieving fast and reliable deliveries with a robust TOC solution simple enough to be supported by standard ERP systems. Sun City, South Africa: SAPICS Conference Proceedings (July 12-15).
- Collins, J. 2006. *Why Business Thinking Is Not the Answer. Good to Great and the Social Sectors: A Monograph to Accompany Good to Great*. New York: HarperCollins.
- Deutschman, A. 2007. *Change or Die: The Three Keys to Change at Work and in Life*. New York: Harper Business.
- Funcke-Bartz, M. 2006. "Sensible constraint management," *Academic Press Plant Systematics*. 78-79.
- Goldratt, E. M. 1984. *The Goal*. Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 1990. *What Is This Thing Called Theory of Constraints and How Should It Be Implemented?* Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 1999a. *Goldratt Satellite Program Sessions 1-8*. (Video series: 8 DVDs) Broadcast from Brummen, The Netherlands: Goldratt Satellite Program.
- Goldratt, E. M. 1999b. POOGI Forum Letters, www.toc-goldratt.com/index.phpcont=137.
- Goldratt, E. M. 2008. *The Choice*. Great Barrington, MA: North River Press.

Goldratt, E. M., Goldratt, R. A., and Abramov, E. 2002. "Strategy & tactics," unpublished article.

Immelman, R. 2003. *Great Boss, Dead Boss*. Gurnee, IL: Stewart Philip International.

Kendall G. 2001. "TOC strategy: The 4 x 4 way," *TOC Review* April:28-31.

About the Authors

Dr. Alan Barnard is the CEO of Goldratt Research Labs, a director of Goldratt Group (Africa), Chairman of Realization Africa and acts as project auditor on a number of large TOC projects around the world. One of Alan's passions in life is to apply the same techniques of identifying and unlocking inherent potential within businesses to individuals (he is the lead facilitators of the annual Odyssey Program) as well as to not-for-profit organizations and governments to help them achieve more with less in less time. Alan is also past chairman of TOCICO (2003-2005), Past President of SAPICS (1998-2002), Chairman of the TOC Odyssey Institute and has been a serving judge in the National Logistics Achiever awards for Southern Africa for the last 7 years. Alan obtained his BSc Industrial Engineering degree (Cum Laude) in 1991 and completed his MSc and PhD in the management of technology and information at the Da Vinci Institute of Technology in 2009. The title of his PhD thesis was "*How to identify and unlock inherent potential within organizations and individuals using the system's approach of Theory of Constraints*".

SECTION V

Strategy, Marketing, and Sales

[CHAPTER 17](#)

[Traditional Strategy Models and Theory of Constraints](#)

[CHAPTER 18](#)

[Theory of Constraints Strategy](#)

[CHAPTER 19](#)

[Strategy](#)

[CHAPTER 20](#)

[The Layers of Resistance—The Buy-in Process According to TOC](#)

[CHAPTER 21](#)

[Less Is More—The Buy-In Process According to TOC](#)

[CHAPTER 22](#)

[Mafia Offers: Dealing With a Market Constraint](#)

Where are we going and how do we get there? This is the classic question of Strategy. In this section we see how to integrate Strategy with Tactics with a view toward continuous improvement. System constraints are a point of focus in the process. Approaches to implementation, the layers of resistance to change that may be encountered and ways of overcoming them are also covered. The application of the basic concepts of Goldratt's Drum-Buffer-Rope and Buffer Management are applied to the management of the

Sales Funnel to increase an organization's sales closing rate and Throughput. The TOC approach, the Mafia Offer, is presented: the guidelines for creating an offer, how to present an offer, and how the reader can create their own.

Two successful approaches to strategy development are provided. Both use the Thinking Processes. In the first approach, Boyd's OODA loop (observe, orient, decide, and act) is applied to the Thinking Processes to provide a proactive approach to strategy. We also see a new paradigm for strategy: the Strategy and Tactic tree. Developed by Dr. Goldratt, this concept links the discrete steps of strategy to the specific tactic each required for success, linking strategy and tactic with assumptions defining their necessary and sufficient interdependence. Strategy and Tactic steps are defined level-by-level to the depth necessary for full and effective execution. Generic strategies are presented here and later in the book outlining recommended approaches for a variety of business environments. "How-to" steps for constructing the Strategy and Tactic tree are described in the next section in [Chapter 25](#).

CHAPTER 17

Traditional Strategy Models and Theory of Constraints

Marjorie J. Cooper

Introduction

Strategy is a difficult concept, and often ill-defined, but the term is used frequently and individuals seem to know at least what *they* consider a strategy. For example, people have strategies for getting things done, from a simple strategy to avoid heavy traffic on their route from work to home to a complicated strategy for acquiring and saving the funds to pay for their children's college education. Some people have strategies for avoiding things, such as putting in too many hours on the job, having to listen to a long-winded neighbor, or being caught by one's spouse while raiding the Godiva chocolates. At the simplest level, strategy is no stranger to most people.

Organizations also have strategies. The processes of developing and implementing these strategies may be detailed and rigorous; often however, strategy development—not to mention the subsequent implementation of the strategy—lacks effectiveness and produces less-than-hoped-for results. Although the field of business strategy within the academic disciplines is comparatively new, theories of strategy conception and implementation have been emerging in business schools and among practitioners for several decades. This chapter will review the major schools

of thought on strategy, including some strengths and weaknesses of those strategy theories. General business strategy, marketing strategy, and sales strategy will each be addressed individually, although an argument will be made that separate strategies attached to specific functions are at best ineffective and, more likely, deleterious to the long-term overriding goals of the organization.

What Is a Business Strategy?

According to Drucker (1994), a *theory of business* is comprised of three parts. These include (1) *assumptions about the environment* in which the business operates or intends to operate, including markets, customers, technology, and societal structures in general; (2) *assumptions about the specific mission* that the company has undertaken; and (3) *assumptions about the core competencies* needed to execute the company's mission successfully. Underlying these assumptions is the notion of integrative ability that allows a company to bring many and widely disparate pieces of information and skills together to form a cohesive implementation plan. Thus, strategy for Drucker is implicitly—if not explicitly—an ongoing process, continually revisited, refined, and updated to address each of these three parts.

Copyright © 2010 by Marjorie J. Cooper.

To define further the foundation for strategy construction and execution, Drucker (1994) mentions four “specifications” regarding these assumptions that must also be true if the firm is to actualize and sustain its objectives. First, assumptions about the company's environment, its mission, and its core competencies must *mirror reality*. Ambitious executives commonly aspire beyond the organization's capabilities or envision a market opportunity

that merely exists in their imagination. For example, in a stunning case of self-deception and institutional hubris, and at the height of American enthusiasm for Japanese-made automobiles, General Motors introduced the Italian-made Allante in 1987 with a price tag of \$56,000+ (Schlesinger, 1988). Clearly, the company failed to appreciate fully the market's resistance toward paying a premium price for a car that was poorly engineered compared to its competitors and perceived by consumers as overpriced. Reality checks are essential.

Second, assumptions in all implementation areas must fit together to form *a cohesive view of reality*, one that offers opportunities that can be addressed by the business. When Howard Schultz, then director of marketing and retail sales for a small Seattle coffee roaster, was strolling through the piazzas of Italy now several decades ago, he was hit with an epiphany (Cuneo, 1994). His vision involved coffeehouses where customers lined up for designer concoctions, paid premium prices, and treated the coffee house as a destination venue. This epiphany became Starbucks, one of the great American brands and a successful business strategy because Schultz's assumptions about marketplace reality were valid.

Third, the theory of the business must be *understood throughout the organization*. That is, members of the organization must be able to articulate and act based on goals, structures, and activities of their specific business: what it stands for and how it operates. Then, it becomes possible for members of the organization to perform the fourth specification, which is to *test their theory of the business against environmental reality*, incorporating the ability to change as needed.

Factors That Comprise Strategy

In a formal sense, a strategy is an integrative plan encompassing the *organization's goals* (or objectives), its *policies*, and its *programs* (Quinn, 1981). *Goals* give focus to an organization's efforts, delineating what major accomplishments the company expects to achieve. *Policies* are the rules and guidelines, both formal and informal, that an organization uses to set limits on the actions of its members. *Programs* are closely linked to the strategic implementation process, providing specified action sequences that direct organizational members toward the achievement of the organization's objectives.

Formalized strategic planning, generally characterized by Drucker's assumptions and Quinn's factors (or equivalents), is highly representative of the process taught in most business schools. However, Quinn's research suggests that the academic ideal is far from common practice. For example, Quinn notes that "Planning activities in major enterprises often [become] bureaucratized, rigid, and costly paper-shuffling exercises divorced from actual decision processes." Moreover, ". . . formal planning often [becomes] just another aspect of controllership—and another weapon in organizational politics" (Quinn, 1981, 42). Most important strategic decisions in the companies that Quinn researched occurred outside of the formal strategic planning process, such that by the end of his research, Quinn noted that ". . . various purported 'normative' approaches to planning began to appear highly questionable, if not actively destructive, in many instances" (Quinn, 1981, 42).

Experience subsequent to the publication of Quinn's research has failed to contradict his findings. In fact, anecdotal evidence throughout the business community suggests that in most cases, the formal "strategic plan" for the coming year (or decade) quickly becomes an exercise in futility and a great waster of human resources.

Criteria for a Good Strategy

In spite of the weaknesses of the strategic planning process as commonly practiced, businesspeople and academics alike, as well as organizational managers from nonprofit and governmental entities, virtually unanimously acknowledge the need for order and some sort of strategic development. Perhaps it would be useful to begin with expectations: What criteria do people expect as outputs for a good strategy? Several essential criteria emerge from the literature.

The two most widely cited *criteria* are *superior performance* and *competitive advantage* (Hill and Jones, 2007, 3). Examples encompassed by these two broad categories include profitability, responsiveness to customers, responsiveness to employees, consistency in meeting objectives, a business model that is customized to the organization's strengths, sustainability, and the ability to handle uncertainty. Thus, contemporary theories of business strategy were developed to focus on these critical outcomes, without which the viability of an organization is in question.

Profitability in TOC terms means that *Throughput*, which is defined *in the TOCICO Dictionary* (Sullivan et al., 2007, 47) as "*the rate at which the system generates 'goal units'.*" A company that does not make money will not stay in business for long (barring government bailouts, of course).

Responsiveness to customers is one of the key requirements for a viable business strategy. Customers vote with their financial resources. If one organization does not recognize and respond to a customer's needs, she will move on to an organization that does. If a company loses enough of its customers, it's out of business.

Since the marketplace is highly competitive, it makes sense for businesses to invest their resources to further develop and support competencies and skills for which the company is already highly capable. Companies that attempt

to take on challenges for which they are underprepared compared to their competitors often experience failure.

Sustainability certainly has environmental connotations today. Companies must be responsive to governmental, regulatory, and community concerns about low impact on the environment. However, sustainability has another important implication for business as well. A sustainable strategy is one that will be successful over time, rather than being a short-term fad that quickly becomes outdated and irrelevant to the marketplace. By leveraging their resources, certain Japanese companies such as Toyota and Canon have made, over time, huge inroads on a global scale, thus enhancing their overall sustainability (Hamel and Prahalad, 1989, 64).

Finally, companies must incorporate into their strategies the means by which they can be *flexible* to changing demands. Product life cycles are growing shorter; consumer preferences adapt rapidly; and economic unrest is pervasive. Thus, companies who are flexible, nimble, and creative stand a much better chance of success than companies that refuse or are unable to change their strategies to meet changing conditions. Part of a good strategy is the built-in capability to respond quickly when a challenge arises.

Theories of Business Strategy

In this section, several theories of strategic management will be examined. Four dominant theories include Ansoff's (1965) matrix of four strategies, Porter's list of strategic foci, learning/emergent strategies, and the resource-based view. A brief review of Mintzberg and Lampel's (1999) summarization of the various schools of strategic management follows. This section concludes with a discussion of the scope of strategic management, which

demonstrates how extensive and demanding the formulation and implementation of a strategic plan can be. The complexity of strategic management perhaps explains in some small way why business strategies often go awry.

Ansoff's Matrix of Four Strategies

In his book *Corporate Strategy* (1965), Ansoff introduced his concept of strategy. He believed that firms should develop a “common thread” that suggests plausible extensions of the firm’s product-market position. Ansoff dubbed these extensions *growth vectors* as represented in the 2 × 2 matrix shown in [Fig. 17-1](#). The four strategies are market penetration, product development, market development, and diversification.

In *market penetration*, the strategy is to increase market share for current products in the markets in which they are currently offered. In today’s competitive marketplace, market penetration generally means not so much finding new users as taking users away from existing competitors—always a challenging strategy.

Product development is the process of conceiving, engineering, and constructing new product solutions for customer problems, such as the development of the flash drive to make possible the convenience of portable data.

Market development is new “missions” for a company’s products. Ansoff really means that new product uses and applications are discovered and promoted to existing customers. The use of Arm & Hammer Baking Soda® as a refrigerator deodorizer is a classic example.

Finally, *diversification*, a more risky approach, moves the company out of its comfort zone into new kinds of products and markets with which the firm has little experience. Diversification adds new businesses to a company not found in its core industry (Hill and Jones, 2007, 340). These sorts

of strategy “jumps” can be immensely profitable but also have a higher-than-average probability of failure because the firm’s competencies may or may not be adequate to make such a leap. Companies often see an opportunity in another industry, for example, that seems attractive only because the company does not fully understand the challenges of entering that industry.

Although Ansoff’s matrix is quite simple and, as viewed today, rather primitive, it continues to be representative of the broad strategic moves open to companies that wish to grow. An important weakness, however, is that such a matrix is purely descriptive rather than prescriptive. It does not tell a company which of several possible growth vectors would be most profitable, most risky, or most easily implemented. Further, the matrix does not provide any sort of plan for the implementation of any growth vectors that may be chosen.

Porter’s List

Ansoff’s list was helpful but hardly comprehensive. His focus was in extending the existing strategic direction of the company. In 1980, Michael Porter introduced a different list of generic strategies, which focused on the initial identification of a business strategy (Porter, 1980). Porter theorized that only two basic strategies were available for companies to pursue—low-cost or differentiation. Porter combines these two strategies in the ubiquitous 2 × 2 matrix format to create the following four options determined by the scope of the firm’s operations: cost leadership industry-wide, differentiation industry-wide, cost focus within a particular segment, and differentiation focus within a particular segment.

Mission \ Product	Present	New
Present	Market Penetration	Product Development
New	Market Development	Diversification

FIGURE 17.1 Ansoff's growth matrix. (From Ansoff, H. I., 1965. *Corporate Strategy*. New York: McGraw-Hill. Original source: Table 6.1, 109. Used by permission of the Ansoff Family Trust.)

Low-cost leadership is generally recognized as a difficult position to maintain in the marketplace. A firm must have a means for achieving sustainable cost advantage. Michael Dell found it in the personal computer industry by putting together logistical alliances with suppliers and an unprecedented Internet-based distribution channel. Other companies have chosen to produce no-frills products, locate a source of cheaper raw materials, improve the overall efficiency of production, or otherwise reduce overhead, such as outsourcing some functions. However, a low-cost strategy is challenging. Even Walmart, with its highly sophisticated logistical system, no longer claims to be the "lowest" price in town.

Differentiation, on the other hand, offers nearly limitless possibilities. Kotler (2003, 318-327) identifies five different bases on which meaningful differentiation may take place: product differentiation, services differentiation, personnel differentiation, channel differentiation, and image differentiation.

Product differentiation can be accomplished by changing the form of a product or adding features. The performance quality of the product can be enhanced and durability may be increased. The product can be made more reliable or easier to repair, thus saving replacement costs. Design is

increasingly a factor that adds value to a company's product offerings as, for example, in Italian-designed leather clothing or German-engineered personal appliances. A related notion is style, which is an important aspect of Apple computers, Montblanc pens, and Harley-Davidson motorcycles (Schmitt and Simonson, 1997).

Services differentiation includes ordering ease and accuracy, on-time delivery, timely and accurate installation, customer education with respect to product usage, and follow-up maintenance and repair.

Companies that have very well-trained sales and service personnel may be able to establish *personnel differentiation*. Lands' End customer service representatives are known for their courtesy and their willingness to go out of the way to assist customers with size, color, and style recommendations.

Image differentiation relies on how the brand's identity is expressed and whether customers develop a strong identification with the company or the brand. Some brands that have established a strong, positive image include Hershey's Chocolate, Coca-Cola, and Betty Crocker. Non-food brands include Mercedes-Benz, Hallmark, and Crayola.

The Resource-Based View

The resource-based view of strategy is primarily based on the analysis of a firm's strengths and weaknesses. This approach draws from three primary research traditions (Barney, 2007, 127-169). The first area of study is *theories of distinctive competence*. Distinctive competence looks at the quality and decision-making skill of the company's general managers who have a very large impact on the company's performance. A second but related area, from a sociological perspective, is the relationship of institutional

leaders combined with the firm's organization and structure to generate distinctive competencies.

The second area of study comes from David Ricardo's (1817) work in economics focusing on land uses and the concept of *economic rent*, "the payment to an owner of a factor of production in excess of the minimum required to induce that factor into employment" (Barney, 2007, 130). For example, a company makes higher profits than its competitors, who are also selling into the same market, because its factory workers have developed production methods with higher productivity rates than competitors. Finally, incorporated into the resource-based view is the notion of *firm growth* (Penrose, 1959), in which an extended typology of sources of productive growth are enumerated and assessed within the administrative framework that links and coordinates group and individual activities. For example, even within the same industry, one company may have the ability to conceptualize new product ideas and translate those concepts into marketable products much more effectively than its competitors.

Representative of these three areas of study combined, the *VRIO framework* stands for a series of questions that must be answered about a company in order to determine what strengths are at the company's disposal as well as what weaknesses must be taken into account. VRIO stands for *value, rarity, imitability, and organization*. Each represents a question with respect to individual resources and groups of resources available to a company, such as whether these resources render a firm capable of responding to threats and opportunities, whether the firm faces cost advantages or disadvantages, and whether policies and procedures are aligned with the firm's use of its resources.

The questions involved in a resource-based analysis enable the firm's strategic planners to identify areas of opportunity that may require upgrading or additional

resources assigned. These questions encourage strategic planners to revisit the company's policies and procedures to assess whether, as currently stated and implemented, they support the company's direction and marketplace positioning. Questions on the rarity of resources help determine the exclusivity of the firm's resources, while questions of imitability help discover cost advantages that the company can exploit.

Learning/Emergent Strategies

Mintzberg and Waters (1985) conceive of deliberate versus emergent strategies as two extremes on a continuum that represent real-world strategies. A team of managers may sit down for a strategic planning meeting and come up with a strategy using analytic skills and deliberation. But that team may just as easily make some important errors in judgment that, subsequently, redefine the team's entire strategic plan. In the former case, the planning is largely *a priori*, structured, and intentional. In the latter case, the resultant strategy may be serendipitous, devised on the spur of the moment under pressure, and as much a surprise to its creators as it was to anyone else. Thus, learning occurs where what was planned does not proceed as envisioned but the outcomes sometimes are better than the planned strategy. Adjustments are made, and the organization moves forward.

According to Mintzberg and Waters, most people tend to treat strategy simply as "an analytic process for establishing long-range goals and action plans for an organization" (1985, 257). However, Mintzberg and his colleagues (viz. Mintzberg, 1978; Mintzberg and Waters, 1982; Mintzberg and McHugh, 1985) conceive of strategy as more fluid than that. Their view is that strategy represents a consistent pattern developed from a stream of organizational decisions

and actions followed by corrections. Because of many studies, what has emerged in Mintzberg's view is the categorization of 10 identifiable schools of thought with respect to strategy. Any of these approaches can lead to success or failure, depending on the situation, although, more realistically, they each represent one piece of the totality of strategic management (Mintzberg et al., 1998, 1-21).

A Summary of Schools of Strategy

Mintzberg and Lampel (1999) distilled the literature on business strategy into various representative schools of thought. They preface their findings with the caveat that each school contributes one perspective or another on the world of strategy, although none offers a comprehensive picture of the reality. The authors identify 10 different schools of thought on strategic management theory, although there is some overlap between schools, especially as cross-fertilization has occurred. The first three are prescriptive, while the latter schools are largely descriptive in their orientation. These schools include: (1) the Design School, (2) the Planning School, (3) the Positioning School, (4) the Entrepreneurial School, (5) the Cognitive School, (6) the Learning School, (7) the Power School, (8) the Cultural School, (9) the Environmental School, and (10) the Configuration School.

The *Design School*, one of the earliest to emerge, focuses on the ability of senior management to find a workable strategic solution while taking account of internal strengths and weaknesses in the light of external threats and opportunities. From this combination of internal capabilities and constraints coupled with external contingencies, a strategy is fitted or designed to best serve the situation in which the organization finds itself. There is no strict,

particularly analytical, or structured formal process, and the central figure in strategy development is the senior executive officer. Intuition of top management is important to the success of the strategy, which is deliberately kept as simple as possible and focused. Until the 1970s, this was the dominant view of strategic management with lasting influence even today.

Ansoff was an important early influence for the *Planning School*, which is characterized by many of the assumptions of the Design School coupled with a formalized, analytic approach to devising business strategy. This process is intentionally step-by-step, supported by techniques that zero in on meeting specific objectives, budgets, and program details.

The *Positioning School* may be the school that most people think of when they think of management strategy. Given impetus by Porter's work as well as the Boston Consulting Group and the PIMS project,¹ this school traces its history back through military strategy to Sun Tzu circa 400 BC. Advocates of the Positioning School believe that strategy can be reduced to generic positions, each representing a formalized compilation of characteristics and industry situations for which generic strategies may be applied as starting templates. The approach is highly analytical; therefore, it is hardly surprising that TOC—with its analytical disposition—has used this approach in identifying and describing its litany of “marketing solutions” according to situational characteristics of the company's milieu.

The *Entrepreneurial School*, like the Design School, focuses on the top executive, but the emphasis is on the cachet associated with a talented individual's mysterious intuition. Rather than using structured methods of developing strategies or analytical means to devise the “right” approach, these strategists focus on the charisma

and talent of a central forceful leader who can use his or her creativity to the organization's ultimate advantage in the marketplace. Moreover, the leader exercises tight control over the details of an implementation.

The *Cognitive School* is comprised of two important orientations. Largely academic, students of this school examine the cognitive processes that people use to arrive at strategies, processes such as mental mapping and conceptual modeling. This school looks at how people reason and arrive at their individual perspectives. The focus is on cognition as information processing, how people are influenced, and what schemas they use to interpret the world. More recently, another view has emerged that focuses on strategies as creative interpretations of reality rather than the objective rendering of reality.

The *Learning School*, through the work of Lindblom (1959), Quinn, (1980), Mintzberg (1978), Mintzberg and McHugh (1985), and others, sees strategy development and implementation as a less structured and intentional process. Rather, strategies evolve incrementally as learning takes place. Formulation and implementation merge, and the result may be something unexpected and creatively discontinuous.

Whether a strategy develops within a firm or as a reaction to forces external to the firm, the *Power School* views the strategy as an essentially political process. Bargaining, persuasion, and encounters between power dispensers within the organization lead to strategies that are in the power network's best interests. The organization thus moves forward based on which powerful individuals and their priorities prevail.²

Mintzberg and Lampel (1999) view the *Cultural School*, a thin stream of research within the academy, as a mirror image of the Power School. Where participants in the power school jockey for power over others—a divisive process—

participants in the Culture School aim for collaboration, continuity, and cohesion within the organization. Thus, highly discontinuous change is discouraged, being seen as leading to dysfunction within the company.

The *Environmental School* may not be, strictly speaking, a school of strategy theory. The focus of this stream of work is on how companies attempt to make compromises with their environments, maneuvering for advantage and success within the confines of external pressures. Mintzberg and Lampel term this school a “hybrid of the power and cognitive schools (1999, 25).”

Finally, the *Configuration School* encompasses two distinct programs. The academic focus looks at organizational configurations, the circumstances in which these configurations evolve and prosper, and the differences according to specific operating conditions, such as a service organization versus a start-up manufacturing organization (Mintzberg, 1979; 1983). The second focus is more practitioner-oriented and deals with how organizations transform themselves, with particular emphasis on change management and the process of moving from one configurative posture to another.

As Mintzberg and Lampel (1999) point out, these schools often define strategy narrowly, failing to take into account aspects of strategy and factors that impact strategy not represented in their respective models. Thus, companies may adopt a particular approach to strategy development and execution that leaves out important considerations for that company. For example, large companies in mature markets may not fare well following the entrepreneurial school of strategy. It is likely that the situation for a large firm is too complex for one person to have a comprehensive understanding of all factors that contribute to meeting the company’s objectives and, at the same time, all the strategic answers for the company’s ongoing success.

Marketing and Strategy

To understand how marketing and strategy fit together, one must understand something about marketing and its scope. In recent years, often marketing erroneously has been conceived of as primarily promotion (Perrault et al., 2008, 7). Although advertising and promotion are certainly areas of importance to marketers in persuading customers of a product's value, a view of marketing that sees the promotion function—advertising, selling, sales promotion—as the primary thrust of marketing is far, far off the mark.

The official (and latest) definition of marketing found on the American Marketing Association's Web site is, "Marketing is the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large" (American Marketing Association, 2007). This definition, nearly all-inclusive of business activities, reflects the evolution of thinking about marketing that was broad even in 1952 when General Electric issued its annual report and stated the following:

[The marketing concept] introduces the marketing person at the beginning rather than at the end of the production cycle and integrates marketing into each phase of the business. Thus, marketing, through its studies and research will establish for the engineer, the design and manufacturing person, what the customer wants in a given product, what price he or she is willing to pay, and where and when it will be wanted. Marketing will have authority in product planning, production scheduling, and inventory control, as well as sales, distribution, and servicing of the product. (1952, 21)

An even more restrictive conception of marketing as the four Ps—product, price, promotion, and place (distribution)—

which has been around since the early 1960s, demonstrates that the marketing function has long been conceived in terms of new product development and refinement, pricing, logistics, and distribution, as well as advertising, sales, sales promotion, and package design. This was not to suggest that marketing was solely responsible for product design and production, meeting shipping dates, or setting prices. Rather, the idea was that marketing represents the customer to the company as well as the company to the customer and is therefore responsible for the strategic alignment and coordination of *all* the factors that attract customers and preserve customer relationships (Perrault et al., 2008, 35–40; Pride and Ferrell, 2008, 3–7).

In 1969, Kotler and Levy’s seminal article, “Broadening the Concept of Marketing,” appeared in the *Journal of Marketing*. Therein, the authors stated, “Marketing is a pervasive societal activity that goes considerably beyond the selling of toothpaste, soap, and steel” (1969, 10). In short order, marketing experts accepted the notion that political candidates market themselves for votes; Hollywood markets celebrities to sell movie tickets; ordinary people market themselves to prospective employers; charities use marketing activities for fund-raising and soliciting volunteers; and organizations market ideas, such as “fasten your seatbelt” and “don’t drink and drive.” From the 1970s on, applications of marketing expanded to a variety of non-business contexts.

Moreover, in today’s marketplace, the concept of *customer relationship management* (CRM) demands that marketing be involved in technology as well (Greenberg, 2001, 1–43). Every single point of contact between the company and the customer is a marketing opportunity, and technology now plays an important role in facilitating company-customer communication and exchanges.

What Is Marketing Strategy?

In light of the scope of marketing discussed previously, one may well say that there is little difference between business strategy and marketing strategy. Ultimately, much of what an organization does strategically will have direct or indirect impact on the firm's ability to cultivate and sustain customer relationships. A small stream of research has even emerged in marketing called internal marketing to address employee issues. This work does not take the place of the huge body of human resource literature; rather, it emphasizes those issues that affect employees in ways that are then secondarily transmitted to customers (Flipo 1986; Adomaitiene and Slatkeviciene, 2008). For example, unhappy employees are more likely to give customers poor service.

Ideally, therefore, the objective is to integrate marketing perspectives and activities into all aspects of a company's business strategy. The benefits are several. First, the customer, without whom the business cannot exist, has representation whenever strategic decisions are undertaken. Second, companies are more likely to avoid making important marketing mistakes that are costly and detrimental to the company's financial position. Third, marketing gains a better understanding of the capabilities and limitations of other functional areas. Fourth, other functional areas are apprised of market demands and competitive actions that affect the company's overall performance. Many other benefits accrue from cross-functional integration as well.

Sales and Strategy

Sales is a subset of marketing. It is the personal promotion function within the firm's total business strategy and the operational arm of marketing (Perrault et al., 2008, 10). In

many (or most) organizations, sales is treated as a separate function from marketing and is managed separately. This approach, of course, breeds a number of difficulties, not the least of which are the tensions that arise between marketing and sales. Sales should be an important set of implementation activities *within the scope of the company's total business strategy*. Unfortunately, silo effects often render null the benefits a well-managed sales function can bring to an organization.

Challenges for Strategy and Execution

A well-executed strategy is the ideal, but the ideal is often far from the reality. A number of potential problems occur that can derail a good plan, much less a plan that has flaws attached to it from the outset (e.g., Colgate and Danaher, 2000). Further, strategy experts themselves admit that few executives are satisfied with the state of strategy knowledge and capabilities currently available (Campbell and Alexander, 1997). As much as 70 percent of business strategies fail to get implemented (Corboy and O'Corrbui, 1999). Beer and Eisenstat (2000) coined the phrase the "silent killers of strategy implementation and learning" and note:

. . . senior managers get lulled into believing that a well-conceived strategy communicated to the organization equals implementation . . . [T]hey approach change in a narrow, nonsystemic and programmatic manner that does not address root causes. (2000, 29)

In this section, some well-documented reasons for strategy failure are cited, including the two central problems mentioned by Beer and Eisenstat—lack of implementation capabilities and failure to address root causes. The TOC body of knowledge directly addresses these and other

problems associated with successful strategy implementation. Some of the contributions made by TOC that address these problems are noted in the following sections.

Inadequate Planning

Strategic planning can be a tedious process, particularly annoying to those types of managers who prefer a more intuitive style of managing, those found, for example, in the design and entrepreneurial schools of strategy. In addition, a formalized process of planning often holds little allure for more action-oriented managers, such as may be found in manufacturing, or people-oriented individuals, such as salespeople. Certainly, good planning demands time, attention to detail, and good communication as well as a willingness on the part of all employees to identify and solve impediments to implementation, negative effects from possible actions, and current practices that block a successful strategy.

However, research indicates that planning does tend to lead to superior performance (Armstrong, 1982; 1991). Because of extensive research on successful strategic planning, Armstrong (2005) identified a five-step program for corporate planning: (1) determining the firm's long-range objectives, (2) generating alternative strategies, (3) evaluating alternative strategies, (4) monitoring implementation and outcomes, and (5) gaining commitment from those who will be affected by the plan. However, a significant difficulty for companies is not only how to complete these rather obvious steps but how to do so using tools that lead to actionable and realistic outcomes. For more information, see Chapters 18 and 19 in this Handbook.

Although planning does not rule out the need for flexibility and creativity, some may view the planning process as

stifling to innovation. However, planning is an important and useful aspect of strategic management, and has been developed to a very high level by practitioners of TOC. Embedded in the Thinking Processes (Scheinkopf 1999) resides a detailed roadmap for strategic planning. These processes include *a set of logic tools that address the critical questions “What to Change?”, “To What to Change?”, and “How to Cause the Change?”* (Originally presented in Goldratt 1990; See *TOCICO Dictionary* (Sullivan et al., 2007) for the change sequence and each question. (© TOCICO 2007, used by permission, all rights reserved.)

Inability to Analyze the System

The system of interest might be a supply chain, members of an industry, or an entire economy. In fact, the system could even be the global economy. However, for the sake of simplicity, one may assume that a system is a company made of up subsystems, such as functional areas. Because of multiple interdependencies between resources within functional areas and between functional areas as well as the impact of the marketplace and environment, cause-and-effect relationships are multitudinous and complex. Often, decisions made in one part of the organization have serious impact on what happens in other parts of the organization and, eventually, on its customers. When localized objectives, power structures, and policies are pitted against one another as each tries to satisfy its responsibilities in strategy implementation, conflict, chaos, and complications result. To develop and implement a strategy that takes these interlocking relationships into account and intentionally limits the conflicts, one must have a method for analyzing the “what if’s” and understanding the inevitable results of various decisions and moves.

Hamel put his finger on a serious problem with the following: “The dirty little secret of the strategy industry is that it doesn’t have any theory of strategy creation” (quoted in Mintzberg et al., 2005, 5). Standard strategy textbooks are largely composed of long lists of things strategists must take into account (e.g., Barney, 2007; Hill and Jones, 2007), but offer virtually no mechanism for conducting a useful analysis or—as Hamel implies—a mechanism for using the analysis to arrive at a plan. Perhaps this is one reason that well-intentioned but highly analytic approaches have not heretofore met with much success. A *detailed* method to help identify potential setbacks as well as the root causes for some knotty problems have been lacking. For example, SWOT (strengths, weaknesses, opportunities, and threats) analysis, widely used and taught in business schools, has almost no analytical value.

The TOC TP, as explained by Scheinkopf (1999), provides a methodical tool for conducting a thorough analysis of a company’s situation to define the root causes of its problems, verify that the hypothesized root causes are, in fact, the true causes, and complete a plan for changing the direction of the company. Included in these tools are methods to identify new marketing opportunities and to apply rigorous analysis to the proposed solution. Methods to promote problem solving and better communication are also included, although it must be admitted that the administrator’s skill in using these tools is a critical factor in their success.

No Theory of Implementation

One of the most ignored yet crucial aspects of strategy is the implementation process. In fact, “Execution is *the* great unaddressed issue in the business world today. Its absence is the single biggest obstacle to success and the cause of

most of the disappointments that are mistakenly attributed to other causes” (Bossidy and Charan, 2002, 5). Put simply, the management discipline has no robust theory of implementation.

A theory of implementation would necessarily require (1) policy and procedure alignments that encourage each subunit to act in ways that support the overall business strategy; (2) performance assessments and reward systems that, likewise, encourage behaviors that are consistent with the overall business strategy; (3) methods of communication that promote cross-functional coordination of work toward a common objective; (4) real-time, reliable mechanisms to order the tasks and priorities of each individual, each subunit, and each functional area in accordance with the overall objectives of the company’s strategy; and (5) ongoing marketplace research, feedback, and communication loops that provide guidance and course correction for strategic initiatives.

TOC shows more progress in meeting these criteria than any other theory of business strategy currently being advocated and taught. First, TOC formally recognizes the need for these criteria. Second, the TOC applications and tools are designed explicitly to address these issues. For example, the TP (see Chapters 24 through 26, this volume) identify the sources of undesirable effects in the system and uncover ways to resolve these problems, which often occur cross-functionally. The TP also suggests methods of communication to facilitate better managerial control of implementations. The TP also provides a basis for evaluating the ongoing marketplace research and feedback so critical to maintaining a high level of customer responsiveness.

Originally developed by Dr. Eli Goldratt, TOC’s Strategy and Tactic (S&T) trees³ (see Chapters 18 and 25, this volume) offer the sort of detailed implementation instructions necessary to keep a strategy focused and

ensure that everyone involved in the implementation is following a coordinated implementation plan. The S&T trees are also helpful in sequencing necessary events in a strategic implementation. The S&T trees at an appropriate level of detail stipulate which application implementations are critical (e.g., *Drum-Buffer-Rope* [DBR], Critical Chain, etc.), in the order they should be applied, and with details of the operational means of putting them into action.

Conflicts within the System

Although the management literature recognizes several types of intra-organizational conflicts, one of the most disruptive types of conflict, the cross-functional conflict, is very common in organizations. For example, marketing managers perceive great dependence on manufacturing, but manufacturing managers perceive significantly less dependence on marketing (Kahn and Mentzer, 1994). Kahn and Mentzer point out that lack of reciprocity thereby inhibits a collective orientation toward an effort to serve the customer (1994, 117). Further, manufacturing managers perceive themselves as more like marketing managers than vice versa on dimensions of work, customer interaction goals, employees, and reward structures, suggesting additional opportunities for misunderstanding and conflict to erupt.

Resolving conflict means that problems that arise during cross-functional integration, which involves different work units interacting and marshalling work, resources, and mutual assistance (Ruekert and Walker, 1987), are minimized (see [Chapter 33](#), this volume). Yet cross-functional relationships must flourish if companies are to compete successfully in today's very competitive marketplace. In the context of strategic management, cross-functional relationships have been found to be integral to

Porter's (1985) value chain, to internal marketing (Ballantyne, 1977), to maintaining a true market orientation (Kohli and Jaworski, 1990), to achieving seamless customer relationship management (Ryals and Payne, 2001), and to aligning marketing and manufacturing strategies with market conditions (Berry et al., 1999).

The challenges entailed in achieving alignment between marketing and manufacturing (viz. Shapiro, 1977; Crittenden, 1992; Crittenden et al., 1993; Hill et al., 1998; Cooper, 2002), marketing and sales (Massey and Dawes, 2007a; 2007b), marketing and engineering (Shaw et al., 2003), marketing and information systems (Cooper et al., 2008), and marketing and project management (Cooper and Budd, 2007) have been explored throughout the business literature (see also [Chapter 33](#), this volume). TOC's TP help identify the core conflicts that exist as generic conflicts between these functional areas and assist in selecting solutions from the TOC array of applications that address these conflicts. The TP can also assist in identifying and resolving idiosyncratic conflicts for which generic solutions are inappropriate.

Conflicting Standards of Performance

A number of academic articles point out the difficulties that arise from conflicting standards of performance. For example, manufacturing may prefer large production runs to lower their cost of production. Marketing, on the other hand, calls for smaller quantities of a variety of products to satisfy the heterogeneity found in differing consumer preferences (Ghose and Mukhopadhyay, 1993). Conflicting standards of performance are a direct contributor to cross-functional conflict, since, inevitably at one or many functional interfaces, these standards clash with the overriding

objectives of the company's strategic focus⁴ (see also Chapters 13, 14, and 33, this volume).

Dysfunctional Compensation and Reward Policies

In the introduction to Steven Kerr's classic article on reward behavior, he states,

Whether dealing with monkeys, rats, or human beings, it is hardly controversial to state that most organisms seek information concerning what activities are rewarded, and then seek to do (or at least pretend to do) those things, often to the virtual exclusion of activities not rewarded . . . Nevertheless, numerous examples exist of reward systems that are fouled up in that behaviors which are rewarded are those which the rewarder is trying to *discourage*, while the behavior he desires is not being rewarded at all. (1975, 769)

For example, management unintentionally causes inflated project task estimates by relying on on-time task completion as a performance measure. This expectation, in turn, incites project team members both to select unrealistically high task completion time requirements and to extend estimates of completion times. Similarly, shortsighted marketing objectives sometimes drive premature commitments by sales staff to delivery of products and capabilities based on schedules that could not possibly be met by developers and engineers.

Kerr's observation focuses attention on a pervasive problem in strategy implementation: poorly aligned performance evaluation and reward systems. Regardless of the extent to which a strategic initiative is on-target and well-thought-out, to the degree to which the system continues to reward behaviors that are contrary to the desired outcomes of the strategic plan, the plan will be

sabotaged and rendered ineffective. McAdam and Bailie (2002), for example, found that not only were performance measures linked to strategy more effective, but also the alignment between the measures and strategy must be continually reviewed and treated as a dynamic and complex issue. TOC not only suggests some measurements that are more directly tied to desirable outcomes in various applications, for example, DBR and *Critical Chain Project Management* (CCPM), but also the TP, used creatively, can suggest customized measures for specific localized and corporate situations. See [Chapter 14](#) of this Handbook.

TOC Contributions

Throughout this Handbook, more detail is provided concerning specific TOC applications and how they contribute to the effectiveness of a firm's strategy. DBR establishes order out of chaos in production, serving to bring reliability, order accuracy, and predictability to the company's production outputs. CCPM not only shows managers how to achieve the same order and dependability in project delivery that DBR has brought to the production floor, but also CCPM and S&T trees are invaluable tools for the implementation of the overall strategy, itself a project or series of projects. The *unrefuseable offer* (URO), sometimes referred to as the "Mafia Offer," is designed to uncover unique, creative marketplace offers that companies are unlikely to consider under normal strategic planning methods. The buy-in process addresses the need to present the URO to the customer or prospect in a manner that is believable and compelling. Sales force engineering demonstrates how an application of DBR to sales force deployment gives a boost to Throughput of a magnitude not usually associated with sales force training and management. TP and S&T trees have already been

discussed in context. The synergy in strategic management generated by the knowledge and *proper utilization* of these tools is unprecedented in the strategic management literature.⁵ Chapters 12 and 16 (this volume) present case studies of holistic implementations of TOC strategies.

Future Research Opportunities

The development of TOC since the initial publication of *The Goal* (Goldratt and Cox, 1984) represents significant progress toward a comprehensive theory of business strategy creation and implementation. Many others have also added to the TOC body of knowledge since that time, many of whom authored chapters for this Handbook. In particular, for more perspectives on TOC and strategy, one might view Chapters 18 and 19 by Kendall and Dettmer, respectively.

Nevertheless, since strategic management still lacks a comprehensive theory of strategy creation, and since management similarly lacks a theory of strategic implementation, opportunities for future research abound, especially if one uses the TOC body of knowledge as a theoretical underpinning for future work. One area of particularly significant challenge involves cross-functional integration. Although the literature abounds with evidence of conflict, the solutions are shallow and have not proven effective. Another area is the application of sales force engineering principles to actual sales force deployment (see [Chapter 22](#); also, Klapholz and Klarman, 2004; Roff-Marsh, 2005). Although literature on team selling exists, the recognition that principles applied in the manufacturing arena could be successful in managing sales forces is yet to be discovered. The use of CCPM in the development and implementation of business strategy is not well known, nor, with a few notable exceptions, is it well documented in the

academic literature. The question of performance measurement and how to devise appropriate measures that support strategic objectives have been explored in general, but tying those measurements to each functional area has not been done well. These are only a few areas in the strategic management field that could benefit from experimentation and reporting of TOC applications and methods.

References

Adomaitiene, R. and Slatkevičienė, G. 2008. "Employee satisfaction and service quality in contact centres," *Economics & Management* 157-158.

American Marketing Association. 2007. Official definition of marketing. <http://www.marketingpower.com/Community/ARC/Pages/Additional/Definition/default.aspx?sq=definition+of+marketing>.

Ansoff, H. I. 1965. *Corporate Strategy*. New York: McGraw-Hill.

Armstrong, J. S. 1982. "The value of formal planning for strategic decisions: Review of empirical research," *Strategic Management Journal* 3:197-211.

Armstrong, J. S. 1991. "Strategic planning improves manufacturing performance," *Long Range Planning* 24:127-129.

Armstrong, J. S. 2005. "Does formal strategic planning really help?" <http://marketing.wharton.upenn.edu/ideas/pdf/Armstro>

ng/ELMAR/Formal%20Strategy%20Planning.pdf
(accessed July 17, 2009).

Ballantyne, D. 1997. "Internal networks for internal marketing," *Journal of Marketing Management* 13:343-366.

Barney, J. B. 2007. *Gaining and Sustaining Competitive Advantage*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Beer, M. and Eisenstat, R. A. 2000. "The silent killers of strategy implementation and learning," *Sloan Management Review* 41(4):29-40.

Berry, W. L., Hill, T., and Klompmaker, J. E. 1999. "Aligning marketing and manufacturing strategies with the market," *International Journal of Production Research* 37:3599-3618.

Bossidy, L. and Charan, R. 2002. *Execution: The Discipline of Getting Things Done*. New York: Crown Publishing Company.

Bower, J. L. and Doz, Y. 1979. "Strategy formulation: A social and political process." In Schendel, D. E. and Hofer, C. W., eds., *Strategic Management*. Boston: Little, Brown, 152-166.

Brandenburger, A. M. and Nalebuff, B. J. 1995. "The right game: Use game theory to shape strategy," *Harvard Business Review* 73(4):57-81.

Campbell, A. and Alexander, M. 1997. "What's wrong with strategy?" *Harvard Business Review* 75(6):42-51.

- Colgate, M. R. and Danaher, P. J. 2000. "Implementing a customer relationship strategy: The asymmetric impact of poor versus excellent execution," *Journal of the Academy of Marketing Science* 28(3):375-387.
- Cooper, M. J. 2002. "Addressing the conflicts between production & marketing," *The Journal of Enterprise Resource Management* 9:23-32.
- Cooper, M. J. and Budd, C. S. 2007. "Tying the pieces together: A normative framework for integrating sales and project operations," *Industrial Marketing Management* 36:173-182.
- Cooper, M. J., Gwin, C. F., and Wakefield, K. L. 2008. "Cross-functional interface and disruption in CRM project: Is marketing from Venus and information systems from Mars?" *Journal of Business Research* 61:292-299.
- Corboy, M. and O'Corrbui, D. 1999. "The seven deadly sins of strategy," *Managerial Accounting* 77:29-30.
- Crittenden, V. L. 1992. "Close the marketing/manufacturing gap," *Sloan Management Review* 33:41-52.
- Crittenden, V. L., Gardiner, L. R., and Stam, A. 1993. "Reducing conflict between marketing and manufacturing," *Industrial Marketing Management* 22:299-309.
- Cuneo, A. Z. 1994. "Starbucks' word-of-mouth wonder," *Advertising Age* March 7:12.
- Drucker, P. F. 1994. "The theory of the business," *Harvard Business Review* September- October:95-

104.

Flipo, J. P. 1986. "Service firms: Interdependence of external and internal marketing strategies," *European Journal of Marketing* 20(8):5-14.

General Electric Corporation. 1952. *General Electric Sixty-First Annual Report*. Schenectady, NY: General Electric Company.

Ghose, S. and Mukhopadhyay, S. K. 1993. "Quality as the interface between manufacturing and marketing: A conceptual model and an empirical study," *Management International Review* 33(1):39-52.

Goldratt, E. and Cox, J. 1984. *The Goal*. Croton-on-Hudson, NY: North River Press.

Greenberg, P. 2001. *CRM at the Speed of Light*. New York: Osborne McGraw-Hill.

Guth, W. D. and Macmillan, I. C. 1986. "Strategy implementation versus middle management self-interest," *Strategic Management Journal* 7(4):313-327.

Hamel, G., Doz, Y. L., and Prahalad, C. K. 1989. "Collaborate with your competitors and win," *Harvard Business Review* 67(1):133-139.

Hamel, G. and Prahalad, C. K. 1989. "Strategic intent," *Harvard Business Review* 67(3):63-78.

Hill, C. W. L. and Jones, G. R. 2007. *Strategic Management: An Integrated Approach*. Boston: Houghton Mifflin Company.

- Hill, T. J., Menda, R., and Dilts, D. M. 1998. "Using product profiling to illustrate manufacturing-marketing misalignment," *Interfaces* 28:47-63.
- Kahn, K. B. and Mentzer, J. T. 1994. "Norms that distinguish between marketing and manufacturing." *Journal of Business Research* 30:111-118.
- Kerr, S. 1975. "On the folly of rewarding A, while hoping for B," *Academy of Management Journal* 18:769-783.
- Klapholz, R. and Klarman, A. 2004. *The Cash Machine*. Croton-on-Hudson, NY: North River Press.
- Kohli, A. K. and Jaworski, B. J. 1990. "Marketing orientation: The construct, research propositions, and managerial implications," *Journal of Marketing* 54:1-18.
- Kotler, P. 2003. *Marketing Management*. Upper Saddle River, NJ: Prentice-Hall.
- Kotler, P. and Levy, S. J. 1969. "Broadening the concept of marketing." *Journal of Marketing* 33:10-15.
- Lindblom, C. E. 1959. "The science of muddling through." *Public Administration Review* 19:79-88.
- Macmillan, I. C. 1978. *Strategy Formulation: Political Concepts*. St. Paul, MN: West.
- Massey, G. R. and Dawes, P. L. 2007a. "Personal characteristics, trust, conflict, and effectiveness in marketing/sales working relationships," *European Journal of Marketing* 41:1117-1145.

- Massey, G. R. and Dawes, P. L. 2007b. "The antecedents and consequence of functional and dysfunctional conflict between marketing managers and sales managers," *Industrial Marketing Management* 36:1118-1129.
- McAdam, R. and Bailie, B. 2002. "Business performance measures and alignment impact on strategy," *International Journal of Operation & Production Management* 22:972-996.
- Mintzberg, H. 1978. "Patterns in strategy formation," *Management Science* 34:934-948. Mintzberg, H. 1979. *The Structuring of Organizations: A Synthesis of the Research*. Englewood Cliffs, NJ: Prentice-Hall.
- Mintzberg, H. 1983. *Power In and Around Organizations*. Englewood Cliffs, NJ: Prentice-Hall.
- Mintzberg, H., Ahlstrand, B., and Lampel, J. 1998. *Strategy Safari*. New York: Free Press.
- Mintzberg, H., Ahlstrand, B., and Lampel, J. 2005. *Strategy Bites Back*. Upper Saddle River, NJ: Pearson Prentice Hall.
- Mintzberg, H. and Lampel, J. 1999. "Reflecting on the strategy process," *Sloan Management Review* 40:21-30.
- Mintzberg, H. and McHugh, A. 1985. "Strategy formation in adhocracy," *Administrative Science Quarterly* 30:160-197.
- Mintzberg, H. and Waters, J. A. 1982. "Tracking strategy in an entrepreneurial firm," *Academy of Management Journal* 25:465-499.

- Mintzberg, H. and Waters, J. A. 1985. "Researching the formation of strategies: The history of Canadian Lady, 1939-1976." In Lamb, R., ed. *Strategic Management Frontiers*. Greenwich, CT: JAI Press.
- Penrose, E. T. 1959. *The Theory of the Growth of the Firm*. New York: John Wiley & Sons.
- Perrault Jr., W. D. Cannon, J. P., and McCarthy, E. J. 2008. *Basic Marketing*. 16th ed. Boston: McGraw-Hill Irwin.
- Pettigrew, A. M. 1977. "Strategy formulation as a political process," *International Studies of Management and Organization* 7(2):78-87.
- Porter, M. E. 1980. *Competitive Strategy*. New York: The Free Press.
- Porter, M. E. 1985. *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: The Free Press.
- Pride, W. M. and Ferrell, O. C. 2008. *Marketing*. 14th ed. Boston: Houghton Mifflin Company.
- Quinn, J. B. 1980. *Strategies for Change: Logical Incrementalism*. Homewood, IL: Irwin.
- Quinn, J. B. 1981. "Formulating strategy one step at a time," *Journal of Business Strategy* Winter:42-63.
- Ricardo, D. 1817. *Principles of Political Economy and Taxation*. London: J. Murray.
- Roff-Marsh, J. 2005. *Reengineering the Sales Process*. Brisbane, Australia: Ballistix Pty Ltd.

- Ruekert, R. W. and Walker, O. C. 1987. "Marketing's interaction with other functional units: A conceptual framework and empirical evidence," *Journal of Marketing* 51:1-19.
- Ryals, L. and Payne, A. 2001. "Customer relationship management in financial services: Towards information-enabled relationship marketing," *Journal of Strategic Marketing* 9:3-27.
- Sarrazin, J. 1977-78. "Decentralized planning in a large French company: An interpretive study." *International Studies of Management and Organization* 7(3-4):37-59.
- Scheinkopf, L. J. 1999. *Thinking for a Change*. Boca Raton, FL: St. Lucie Press.
- Schlesinger, J. M. 1988. "GM halts output of Allante model; sales link denied." *The Wall Street Journal* June 10:1.
- Schmitt, B. and Simonson, A. 1997. *Marketing Aesthetics: The Strategic Management of Brand, Identity, and Image*. New York: The Free Press.
- Shaw, V., Shaw, C. T., and Enke, M. 2003. "Conflict between engineers and marketers: The experience of German engineers," *Industrial Marketing Management* 32:489-499.
- Shapiro, B. P. 1977. "Can marketing and manufacturing coexist?" *Harvard Business Review* (September-October):104-114.
- Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>

About the Author

Marjorie J. Cooper received her PhD from Texas A&M University and spent four years on the faculty at the University of Texas at Arlington. She has been on the faculty at Baylor University since 1984 and has held positions as Research Director and Executive Director for the Center for Professional Selling.

Dr. Cooper is author or coauthor of five books—including *Introduction to Marketing* published by HarperCollins Publishers in several countries—and seven book chapters as well as published online course materials for retailing and advertising. She has published more than 100 articles in trade publications, such as *Association Management*, *Counselor*, *Promotional Products Business*, *Inside Impact*, *Imprint*, *Spiritual Fitness in Business*, *Business Press*, *Sacramento Business Journal*, and others, including a number of cover stories. She has also published 10 business cases, which were used as the bases for national student competitions. Her academic publications include more than 50 journal articles and published papers with placements including the *Journal of Marketing*, *Journal of Advertising*, *Journal of Advertising Research*, *Journal of the Academy of Marketing Science*, *Journal of Personal Selling & Sales Management*, *Industrial Marketing Management*, *Journal of Business Research*, *Business Horizons*, *Journal of Promotion Management*, *Journal of Marketing Education*, *Journal of Hospital Marketing*, *Journal of Business & Industrial Marketing*, *Psychology & Marketing*, and others. Dr. Cooper's work has received "Best Paper" awards from the *Journal of Marketing Education* and the Southwestern Marketing Association.

Her current research interests focus on customer/company identification and the impact of practices in operations management on a company's ethical

climate and productivity. She has published several papers on the Theory of Constraints, including three on Critical Chain Project Management.

Her consulting and research background includes 26 years of working with the promotional products industry. She is internationally recognized as the foremost academic authority in the promotional products arena and is the principal researcher and a coauthor for *Counselor's* Annual State-of-the-Industry issue. For 17 years, she was also the principal researcher of the sales volume estimate for the entire promotional products industry.

She has an active consulting practice with a variety of clients and previously served on the Board of Advisors for an Internet start-up in Silicon Valley, California. She is a Jonah and has been an Academic Associate of the Goldratt Institute in New Haven, Connecticut. She was also a member of the Institute's Joint Executive MBA Alliance Team, and is a founding member of TOCICO, the international certification body for the Theory of Constraints.

CHAPTER 18

Theory of Constraints Strategy

Gerald Kendall

Introduction—What Differentiates a TOC Strategy?

Over the past 20 years, some remarkable leaders have shared a wealth of knowledge on strategy. If you have read bestsellers such as Bossidy and Charan's (2002) *Execution*, Collins' (2001) *Good to Great*, or Kotter's (1996) *Leading Change*, you may wonder what the difference is in a *Theory of Constraints* (TOC) strategy. To understand it best, consider this example. A Canadian manufacturer had taken several years to achieve a *due-date performance* (DDP) on custom orders of close to 80 percent, better than average for their industry. Within 3 months of implementing a TOC strategy and tactics, their DDP jumped to the high 90s. What made the difference, in one word, is *focus*.

TOC strategy is as much about choosing to *not* do many things as it's about *focusing* on the single biggest leverage point of an organization. TOC assumes that within every complex system, there is "inherent simplicity." This implies that there are very few constraints (leverage points) that determine the performance of a system.

The key to driving success with a TOC strategy is to keep an executive team focused on their single biggest leverage point long enough to sustain results. The most recent TOC strategies are designed to drive exciting, measurable results

within a few weeks. Without such an effect, it's easy for any team to lose energy and readily choose diversions.

This chapter explains the two main formats of strategy that TOC uses to create quick, measurable success. It also describes the generic focus points identified and suitable for many manufacturing, distribution, and project environments. The *Strategy and Tactic Tree* (S&T) knowledge shared in the figures throughout this chapter would not have been possible without the generous agreement of Dr. Eli Goldratt to make this content available throughout the Handbook.

Chapter Overview

TOC's strength in organization strategy is bringing focus with a holistic approach. However, the TOC contribution to the science of strategy goes well beyond that. There has been a profound level of thinking and a masterful set of experiments conducted to bring improvement levels and speed of implementation to new heights. One CFO, who had been part of a venture capital firm responsible for billions of dollars of investments, told me that he had never seen such a comprehensive S&T document for any business.

Copyright © 2010 by Gerald Kendall.

This chapter summarizes the current knowledge, which is applicable to any organization (for profit or not-for-profit, government, etc.) in any industry (manufacturing, distribution, retail, service, etc.). To help you quickly grasp what I have learned in 15 years of applying TOC to strategy and tactics, I have organized this chapter into ten distinct sections:

Introduction—What Differentiates a TOC Strategy?

Chapter Overview

Definitions and Foundations of TOC Strategy

Overview of TOC Strategy Applications in Manufacturing, Projects, and Consumer Goods Distribution/Retail Organizations

Four Generic Prerequisites/“Injections” for a Lasting Competitive Edge

Desirable Effects of a Good Strategy

Two Forms of Strategy and Tactics—TP and S&T Trees

Integrating Other Methodologies, such as Lean and Six Sigma

Dealing with Human Behavior in a Strategy Summary

Definitions and Foundations of TOC Strategy

Strategy defines *what we want to achieve*. *Tactics* describe *how we plan to achieve the strategy*. If the two become separated, it's easy to go astray and fail. TOC strategy comes with proven tools, which clearly link strategy and tactics for all levels in an organization. These tools are the key to success because people resources are scarce and management attention is limited. Simply put, the tools are designed to prevent rework.

There is a great danger in simply stating a strategy as a set of high-level objectives. For example, if the top executive team sets a strategy of becoming the most profitable company in their industry, there are many different directions the company could take to get there. They could focus on higher margin products, mass cost cutting and efficiency programs, expanding markets, or increasing market share. The possibilities of misalignment between functional areas and of making mistakes in implementing a strategy point to a need for much greater clarity and depth. TOC provides clarity through a hierarchy of linked strategies and tactics at multiple levels (see S&T example, [Fig. 18-1](#)) or through a detailed road map of cause and effect between “injections” (ideas to be implemented)

and desired effects (see Future Reality Tree example, [Fig. 18-2](#)). In this chapter, I explain both approaches and the differences.

Three Goals or Necessary Conditions of Any Strategy

For any organization, a good strategy will achieve and sustain three essential effects:

1. Increase goal units (e.g., net profit, education units, health units) now and in the future.
2. Satisfy markets (customers, suppliers, community) now and in the future.
3. Provide employee security and satisfaction now and in the future.

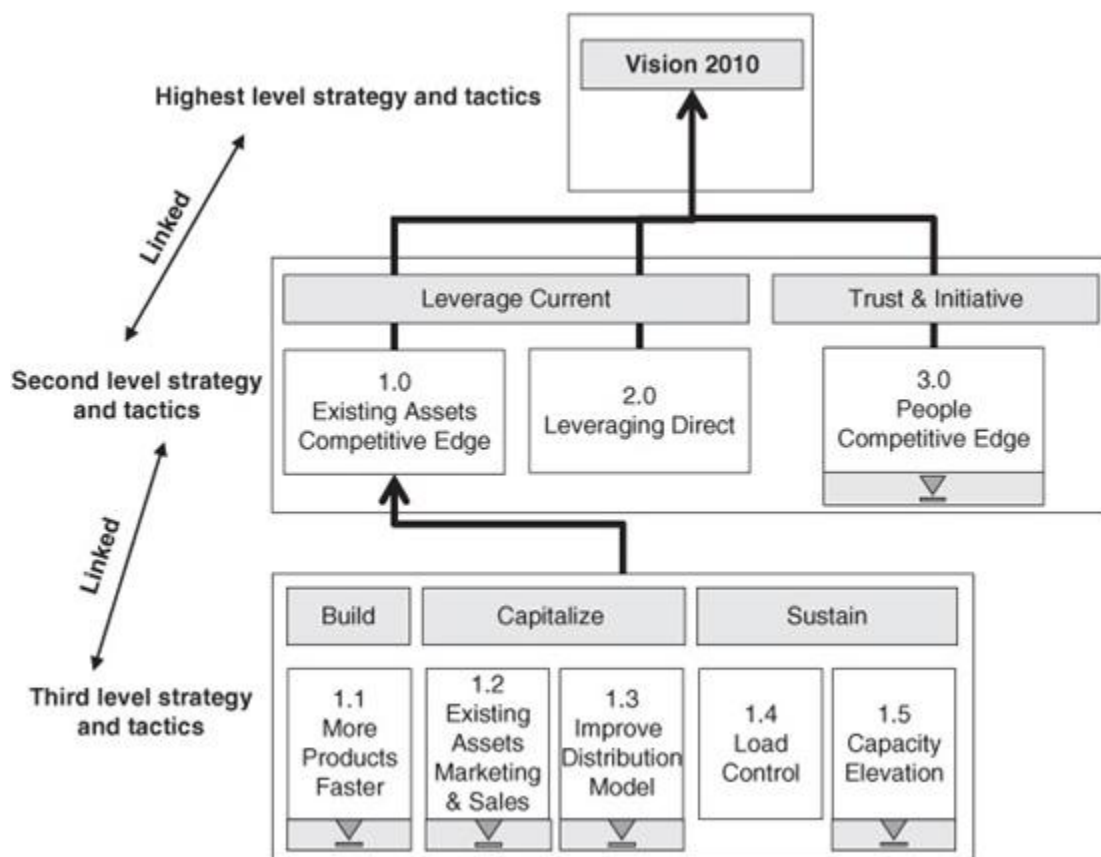


FIGURE 18-1 The Strategy and Tactics Tree. (Used by permission of E. M. Goldratt, S&T © E. M. Goldratt.)

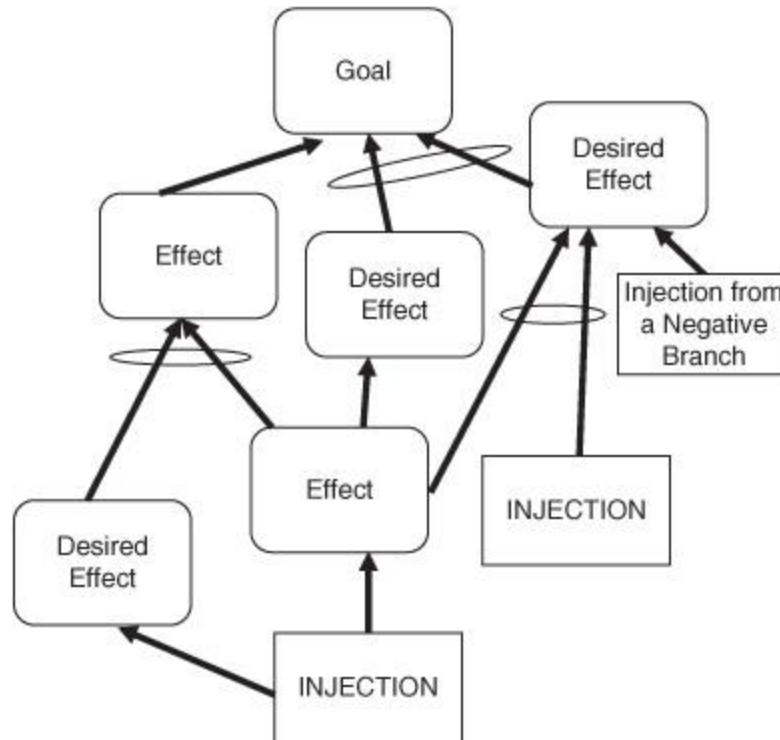


FIGURE 18-2 The Future Reality Tree Thinking Process to define strategy.

You could make any one of these your goal. As long as you consider the other two as necessary conditions, your overall strategy will be the same. This means that a strategy is incomplete if it does not encompass, explicitly, all three necessary conditions. For example, in high growth strategies, organizations often don't cover all the necessary processes to keep the organization stable in terms of customer satisfaction and employee skills and development.

The goal units typically appear at the top of a TOC strategy structure. While it's up to an organization's owners or key stakeholders to set their goal, TOC embeds a higher goal—a *Process of Ongoing Improvement* (commonly referred to as POOGI in TOC)—into its methodology. The

assumption is that every organization has a constraint, always, and therefore can continually improve. With POOGI imbedded into a strategy, improvements are sustainable year after year.

However, even with a well-defined goal, organizations, using traditional processes, often don't achieve it.¹ TOC would suggest either that such organizations focused on the wrong initiative (a non-constraint) or that management attention was constantly diverted away from the right leverage point. One current belief is that the biggest constraint to improving any organization is management attention. TOC is not perfect, but the literature and research shows a higher success rate. In my opinion, this is not by accident, but by the power of a TOC approach called the *Five Focusing Steps* (5FS), which holds management attention in the area needed to deliver quantum results.

The Five Focusing Steps

TOC strategy finds and capitalizes on inherent simplicity in any organization by using a process called 5FS (Goldratt, 1990, [Chapter 1](#)). These steps are:

1. *Identify* the system's constraint.
2. Decide how to *exploit* the system's constraint.
3. *Subordinate* everything else to those decisions.
4. *Elevate* the constraint.
5. *Go back* to Step 1. **Warning:** Don't let inertia become your system's constraint.

For organizations that fit an S&T template (see examples in the next section from several predefined, generic

solutions to common problems), the 5FS have already been applied and are reflected in the S&T. Otherwise, the organization defines its strategy and tactics by answering the implicit questions in the 5FS. If the people defining the strategy are stuck in how to apply one of the 5FSs, one or more of the TOC *Thinking Processes* (TP) can be used to derive an answer.²

The following example illustrates the 5FS. You will find the description of generic S&T strategies later in this chapter and in Chapters 25 and 34.

Example—The Five Focusing Steps

A retailer wants to increase sales and profits. In applying the first step, *identify*, the retailer looks for a physical representation of the constraint. Within a distribution system, there are several physical attributes—for example, the physical display space (shelf space) in a retail store, the capacity of trucks that make deliveries, or the number of parking spaces available to customers. In most retail environments, TOC finds the biggest leverage point (physical constraint) represented by the customers who come to buy the products.

Since the word *exploit*, in English, can have a very negative connotation (i.e., one person taking advantage of another person in a very negative way), you can substitute the words “Get more out of.” Often, when a constraint is identified, upon closer examination there is significant waste of the constraint. In distribution, one of the biggest wastes is attracting customers to a store and not having the product they desire in stock when they arrive. This is the opposite of *exploit*. Simultaneously, there is another huge waste—having too much inventory in some locations. The waste is not just in carrying costs of the inventory. More often, the bigger waste lies in the efforts that go into getting

rid of such inventory, which in today's world can quickly become obsolete, and in the waste of prime retail space occupied by slow moving items. In order to *exploit* the constraint in the retail system, TOC advocates moving from a *push* system (where high quantities of goods are pushed from the manufacturer to a distributor and from the distributor to a retailer through policies rewarding buying high quantities in a single shipment) to a *pull* system (where you replenish more frequently exactly what was consumed, and modify inventory target levels dynamically and automatically according to significant changes in demand).

The *subordinate* step defines the actions and policies required by everyone in the system in order to support the decisions made in the *exploit* step. The most successful implementation of TOC subordination in distribution that I witnessed required store managers to stop deciding on what to order and instead concentrate on training their retail staff in serving their customers (see [Fig. 18-3](#) for actual results from this S&T tree). By subordinating everyone to the TOC pull system ordering decisions, sales doubled within 3 months with the same level of inventory and, *at the same time*, shortages were cut in half. Another way to decide to exploit is to replenish more frequently to drive down shortages, allowing the retailer to carry a greater variety of items—the more frequent the replenishment, the smaller the quantity of individual items needed to meet a given customer service level.

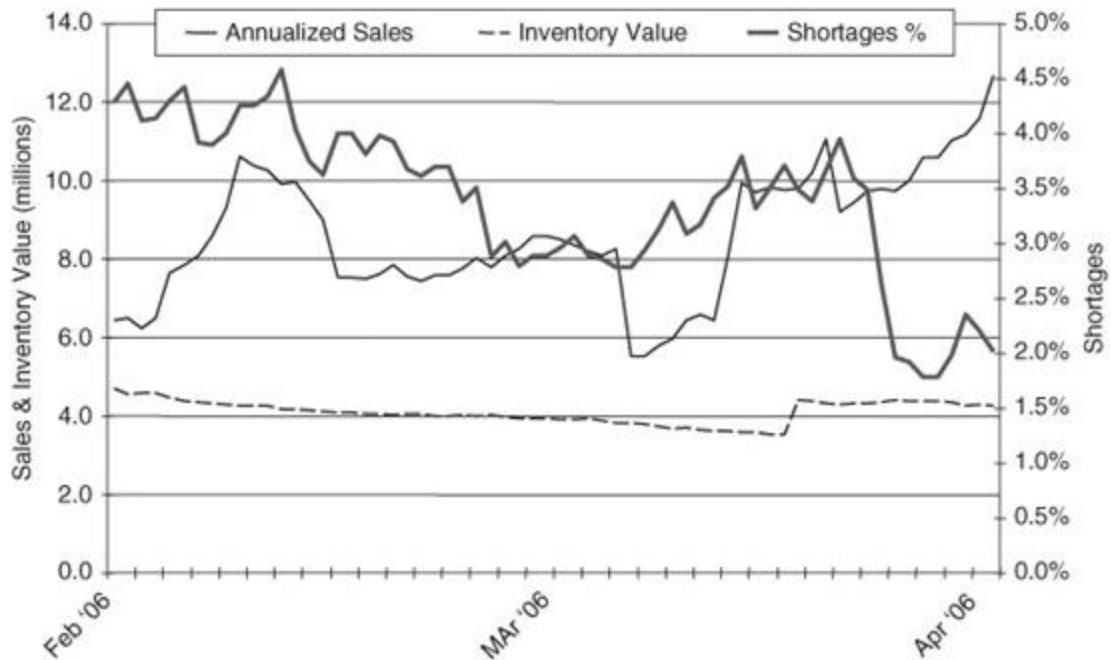


FIGURE 18-3 TOC Five Focusing Steps in retail—results in 3 months.

Elevate means increasing the capacity of a constraint. Most retailers must choose to carry a very small percentage of all the products available to them, due to physical space limitations. This implies that customers will come to buy a product and a specific retailer will often not have the product they want. For example, the distributor I worked with had 200,000 items available to sell across Europe, but their typical retail store only carried, at most, 2000 items (with variations of the same item by color and size). To elevate the constraint, you might carry out extensive advertising, expand the physical capacity of the store, or increase shelf space to increase the number of items carried in order to drive more sales.

It's important to note the sequence implied by using the strategy of the 5FS. Before the *elevate* step, which often involves investing money, the sequence of the steps makes sure that you first at least try to squeeze more out of what you already have.

When the constraint changes, the last of the 5FS asks you to *go back* to Step 1—*identify* the new constraint. Ideally, since there is so much work in progressing through the 5FS, a constraint should seldom change. At least, you should be able to anticipate a change and prepare for it so that your markets are not impacted. However, in my experience, the western world paradigm of the past 20 years of ordering items in large quantities from overseas has caused the retail constraint to move frequently and unexpectedly to other parts of the supply chain. The customers came to buy, but could not get the item they wanted because the retailer could not get replenishment from their supplier. This was not a temporary shortage until the next arrival of a truck to the store, but a more permanent out-of-stock situation in country distribution centers. Through ordering much more than was necessary to satisfy short-term demand, the western world put their overseas manufacturers into a situation of being weeks or months away from replenishing the western world supply chain. The TOC strategic solution is to order much smaller initial quantities and replenish more frequently. The increased shipping costs and potential increases in product costs are usually offset by a factor of several times by the value of increased sales and reduced obsolescence.

The Role of Throughput Accounting and Other Metrics in Strategy

One of the deficiencies I have observed in many strategies is the absence of a tight link to measurable results. Scorekeeping, combined with well thought out actions to impact the score in the short term, are frequently missing from the outset of launching the implementation of a strategy. The TOC method of holistic scorekeeping called *Throughput Accounting* (TA), described in Chapters 13 and 14 of the Handbook, can be used to address this deficiency.

It is vital for the executive team, managers, and employees of a company to be energized by results. While I don't discount the value of intangible results, such as feelings of satisfaction and trust, perceptions are often different in the intangibles, depending on where you sit in an organization. However, I find that most people view the TOC tangible scorekeeping as more believable because it removes the distortions and subjectivity of cost accounting allocations.

TOC TA recognizes improvement in terms of the impact on *Throughput, Investment, and Operating Expense (OE)*, or derivatives of these (such as net profit). These impacts often lag the effort by weeks or sometimes months. Therefore, leading indicators are also used. Some examples of leading indicators relative to net profit, used effectively in TOC strategy, are:

- Manufacturing—DDP percentage, buffer penetration in red zone percentage
- Projects—number of late projects, priorities driven by red tasks, buffer recovery trend
- Distribution/Retail—shortages percentage, average inventory turns, sales per square meter or foot of shelf space
- Sales for any organization—average sales cycle duration, hit ratio on prospecting
- Health care hospital emergency room—patients discharged in 4 hours or less percentage
- Quality—rework daily or weekly OE percentage of sales

Overview of TOC Strategy Applications in Manufacturing, Projects, and Consumer Goods Distribution/Retail Organizations

Introduction to Strategy Applications

If you want to see the result of 25 years of TOC development in strategy and tactics in a single document, this information is now available in the public domain at no charge. Check out the viewer provided by Harmony software that comes with generic *Strategy and Tactic Trees (S&T Trees)* ³ for manufacturing, projects, and distribution/retail of consumer goods: <http://goldrattresearchlabs.com/?q=node/2>

Since the details of these S&Ts are provided in the above forum, this part of the chapter will serve as a roadmap to these strategies.

What information should be contained in a strategy description?

1. There must be enough detail so that anyone, at any level and function in the company, can understand what they must do in order to meet the company goals. This has major implications for anyone tasked with creating a strategy for an organization. Simply listing objectives and actions or injections is only a start to providing the clarity of direction for all levels and functions.
2. There should be an answer to the question, “Why is this result necessary to achieve the goal or the next higher level strategy?” In this knowledge worker age, most people don’t want to be told *what* to do. Their commitment requires them to understand *why* they need to do it.
3. The strategy should spell out what components will be sufficient to achieve the desired intermediate

results and the overall goal. By forcing ourselves to try to understand the full context of what is required to achieve a major goal, we often discover missing pieces a priori. This alone can significantly speed up an improvement process and prevent painful mistakes. Consider an example where a manufacturer improves internal operations to a point of increasing capacity by 30 percent, and the lead time to increase sales to take advantage of the extra capacity is 6 months. A strategy that plans for the sales lead time will achieve results months earlier than one that is insufficient in the sales portion.

Generic Content of S&T Structures

Before a brief tour of three generic strategies for manufacturing, consumer goods distribution, and projects, it's helpful to understand the format of every S&T. An S&T consists of a structure (see Manufacturing S&T example in [Fig. 18-4](#)) and corresponding text containing strategy, tactics, and assumptions (see corresponding text for box 1 [upper left-hand corner of table provides table position in tree] of the Manufacturing S&T example in [Table 18-1](#)). In the Harmony S&T viewer,⁴ double clicking on a box brings up the corresponding details. By double clicking on box 1 of the manufacturing S&T, labeled *Viable Vision* (VV), you would see the format shown in [Table 18-1](#). Clicking on a "+" symbol in the Harmony software opens up the structure to lower levels.

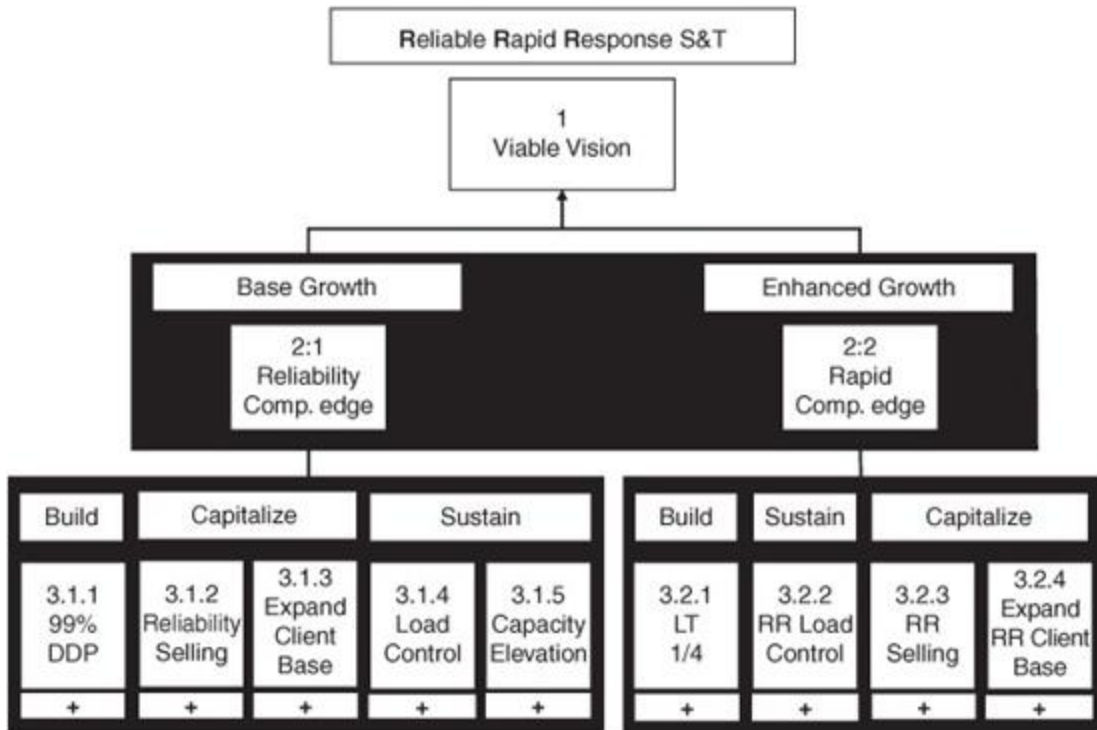


FIGURE 18.4 Manufacturing S & T Structure to three levels. (Used by permission of E. M. Goldratt 2008. *Manufacturing S&T* © E. M. Goldratt.)

The structure is hierarchical. Each lower level in the diagram contains strategies and tactics necessary to achieve strategies represented in the higher level. Ideally, implementing the tactics at the lowest level is sufficient to achieve the strategy at the next level. For example, when the tactics in all boxes below level 2.1 (i.e., 3.1.1, 3.1.2, 3.1.3, 3.1.4, and 3.1.5; use Harmony S&T Viewer to view these tables) have been successfully implemented, the S&T in higher level 2.1 should be achieved.

In every S&T, the strategy is shown in the box labeled *Strategy* as illustrated in [Table 18-1](#). The concept of the VV goal was originated by Dr. Eli Goldratt and first verbalized as turning top line (revenues) to bottom line (net profit) in 4 years or less. Currently, the VV goal is more often expressed as achieving a level of measurable performance beyond

what the management team is targeting. For example, one VV company is a \$400 million company targeting to reach more than \$500 million in two years. That is a stretch for them because they have been on a path of declining revenues for a long time. Just to grow again requires a major change in their thinking and in strategy and tactics.

1	Manufacturing Reliability Competitive Edge
Strategy	(The Company is solidly on P00GI) Viable Vision (VV) is realized in 4 years or less.
Parallel assumptions	For the company to realize the VV its T must grow (and continue to grow) much faster than OE. Exhausting the company's resources and taking too high risks severely endangers the chance of reaching the VV.
Tactic	Build a decisive competitive edge and the capabilities to capitalize on it, on big enough markets without exhausting the company's resources and without taking real risks.
Sufficient assumptions	The way to have a decisive competitive edge is to satisfy a client's significant need to an extent that no significant competitor can. [For different situations, different templates satisfy this condition. The tree below is valid for situations where reliable rapid response template conditions apply to (almost) the entire market of the company.]

(Used by permission of E. M. Goldratt 2008. Manufacturing S&T © E. M. Goldratt.)

TABLE 18-1 Manufacturing S&T Strategy for 1- Reliability Competitive Edge

In a VV S&T Tree, the strategy strives to keep the constraint usually in the market. Internal capacity is managed to be a step ahead of the market. This is accomplished through a combination of a TOC logistics solution and processes to manage capacity. For example, in the manufacturing S&T Tree, the scheduling and execution logistics called *Simplified Drum-Buffer-Rope* ⁵(S-DBR)⁶ (see [Chapter 9](#)) is used in combination with additional steps for controlling promises to the markets when internal

constraints emerge. The S&T Tree also includes steps to elevate capacity proactively in a more timely way.

Parallel assumptions shown in [Table 18-1](#) provide insight into why the tactics make sense as a means to achieve the strategy. The assumptions may indicate facts or warnings you must take into account in coming up with tactics or facts that will make it possible or easier to achieve the strategy. Since VV goals in this strategy imply fast growth in net profit, the parallel assumptions in box 1 contain two warnings:

- First Parallel Assumption warning—Don't try to achieve this vision by cost savings. It must be done by growing Throughput continuously.
- Second Parallel Assumption warning—If you don't get results quickly, you will exhaust one of your most precious resources—the top management team. Do that and you have a high likelihood of killing the entire strategic effort.

Tactics explain *how to achieve the strategy*. At this high level in the S&T tree, the statements are very general conditions that must be implemented, much as you might find expressed in “*Injections*” in a *Future Reality Tree* (FRT; part of the TOC TP). At lower levels in the tree, the tactics represent actions that individuals or small teams will take, much as you might find in a TOC *Transition Tree* (TRT). In the one tactics statement shown in [Table 18-1](#), which is generic for many S&Ts, the organization is told to implement five components:

1. *Build a decisive competitive edge*—How? The answer is detailed at lower levels. What is a *decisive competitive edge*? *It's a capability that the market*

needs to get a significant result. It's so powerful that the company will be able to choose its customers instead of its customers choosing them. The company will be able to grow its business at will. In the Manufacturing S&T strategy, for example, the decisive competitive edge is reliability, as represented by sustained, consistent DDP that is much higher than the industry average. In 3.1.1 (not shown here, use Harmony S&T Viewer), the S&T calls for using S-DBR to achieve the reliability strategy.

The tree goes on to explain in level 2.1 (see [Table 18-2](#)) that such an edge only applies to markets where suppliers are very unreliable combined with the condition that the customers experience significant damage from the unreliability. This is a new type of assumption, called a "*necessary assumption*," found in every box below the first level, which explains why the strategy is necessary or why the strategy takes a given direction.

2. *And the capabilities to capitalize on it*—These are the sales and marketing capabilities. Some of the greatest advances in TOC thinking over the past five years have been in implementing sales practices, lead generation, and processes to increase sales success. The processes define target marketing, selecting "business deal" salespeople, understanding the real benefits of reliability to a customer, creating the offer material, preparing sales people, test launching the offer, defining the sales process steps, fully rolling out the offer with all sales people, monitoring the sales process to continually improve and prevent prospects from being lost, and generating leads so that salespeople can focus on selling. You can use the S&T viewer referenced above to view marketing and sales slides and levels below to understand these

components of the sales and marketing capabilities better.

2:1	Reliability Competitive Edge
Necessary assumptions	When the due dates of the suppliers are notoriously bad and late delivery has major consequences for the client, reliability is a client's significant need.
Strategy	A decisive competitive edge is gained by the market knowing that the company's due-date promises are remarkably reliable, when all other parameters remain the same.
Parallel assumptions	Promises are cheap. Putting money to back up promises (especially when no one else dares to do the same) is convincing.
Tactic	The company is remarkably good at meeting its promised due dates and offers hefty penalties for each time interval of delay. Hefty penalties means enough to deter a competitor from offering (or even from yielding to pressure to do) the same.
Sufficient assumptions	Building a decisive competitive edge is not easy; building the capabilities to capitalize on it is no less difficult. However, sustaining these two elements is the real challenge.

(Used by permission of E. M. Goldratt 2008. Manufacturing S&T © E. M. Goldratt.)

TABLE 18-2 Manufacturing S & T for 2.1 Reliability Competitive Edge

3. *On big enough markets*—When a company's market share is high (e.g., 50 percent or greater) or the target market is too small (e.g., only a handful of prospects exist that meet the market criteria), the chance of reaching an aggressive strategic goal is greatly diminished. The entire S&T is designed to work with much higher odds than those found in Las Vegas.
4. *Without exhausting the company's resources*—Especially without exhausting the top management team. If management attention is truly the constraint

of achieving an aggressive strategy, the way to protect it is with quick results. The lowest level boxes of an S&T should get significant results within the first 2 to 3 months of effort. Significant means measurable increases in Throughput or net profit or in the parameters, such as DDP, which are needed to affect the bottom line.

5. *And without taking real risks*—In this S&T Tree, the lower levels recognize that one of the biggest risks is the risk of success—huge success. Imagine what would happen if you acquire many new clients based on reliability and cannot cope, operationally, with the success. Instead of DDP being sustained to meet customer promises, suddenly the entire system is stretched beyond its limits, reliability collapses or lead times are extended indefinitely, and damaged customers are lost permanently. Such consequences are avoided through processes outlined in lower levels where risks such as internal capacity constraints are identified and monitored. In addition, the S&T Tree includes a capacity elevation proactive process (see S&T section 3.1.5 [use the Harmony S&T Viewer] and below this level for details) to increase capacity more quickly when needed. The tree also contains sales processes to control development and release of prospects to avoid the risk of losing prospects due to lack of attention.

Each of the specific applications below is summarized from S&T designed to achieve a decisive competitive edge.

Manufacturing

[Figure 18-4](#) shows the first three levels of the Manufacturing S&T Tree. [Table 18-1](#) shows the words behind box 1. In this

S&T tree for *make-to-order* (MTO) manufacturers, called *Reliable Rapid Response* or triple R, the strategy is first to achieve a level of reliability unmatched in the industry, as expressed in DDP. In the 1990s, with ample evidence of the ability to improve manufacturing flow quickly through the application of TOC's logistics solution called *Drum-Buffer-Rope* (DBR; see Chapters 8 and 9), the constraint often moved to the market—the manufacturing company was challenged to find enough customers to take advantage of the added capacity quickly enough to avoid pressure to lay off shop floor people. Dr. Goldratt introduced the concept of a "*Mafia Offer*"—*an offer to the market that was too good to refuse*.

One of the great examples of a manufacturer-constructed "*Mafia Offer*" presented at two TOC conferences was California's Orman Grubb furniture company. Jeff Grubb dramatically increased sales of home office and entertainment furniture by making his "*Mafia Offer*" to his retail customers. His offer was to replenish, bi-weekly, any items of furniture bought in individual stores without penalties for freight or penalties for ordering small quantities. In the 1990s, where other manufacturers were forcing retailers to order large quantities per shipment, the offer was well received. The retailers saw that their most popular items would be replenished quickly and that they would not be stuck for months with less popular items. However, by April 2004, overseas manufacturers offered such a huge difference in price that Grubb's offer was no longer sufficient to compete. The company filed [Chapter 11](#) bankruptcy.⁷

Scientists embrace failure as an opportunity to learn, and Orman Grubb teaches a valuable lesson in strategy. The current versions of the S&Ts recognize that an offer to the market is only valid as long as all other factors (price, quality, etc.) remain the same (see [Table 18-2](#), strategy

statement). Either an offer must create sufficient value to overcome price differences or the company must diligently ensure that their strategy contains pieces to keep other factors the same. Every word in an S&T counts—the words come from vast worldwide experience in implementing solutions.

The Manufacturing S&T has two major branches—2.1 (see [Table 18-2](#)), representing the reliability competitive edge and 2.2 (see [Fig. 18-4](#)), representing the rapid response competitive edge. The S&T are generally implemented in a left-to-right sequence. In Branch 3.1.1, all of the basic processes are implemented to achieve an unheard of level of DDP, which is the assumed reliability needed to surpass competitors. Of course, this necessary assumption is not true for every manufacturer in the world. It's true if the manufacturer has large enough target markets of customers suffering from unreliability of significant competitors of the manufacturer. When this has been implemented, we proceed to Reliability Selling as indicated in 3.1.2. (not shown here, use the Harmony S&T Viewer)

When constructing such a strategy, it's hoped that the manufacturer can fully achieve their strategic goals using only the left branch of the tree—2.1 and those entities underneath. The right side beginning with the elements in 2.2 (not shown here, use Harmony S&T Viewer) can be called the bonus or the gravy for achieving or surpassing the strategy. In this tree, the left side proves that the company has the discipline and execution skills to achieve a target that few, if any, in their industry have achieved. The processes, including capacity elevation, marketing, and sales, are stable and provide a base to execute on the right side.

On the right side, using the TOC POOGI process for S-DBR (see [Chapter 9](#)) and *Buffer Management* (BM), production lead time is further reduced so that without taking special actions in manufacturing, the company is capable of filling

custom orders in much shorter quoted lead times to customers. However, the shorter lead times are provided at a high premium to customers where they don't have ready reliable alternatives and where the potential damage to them from regular lead times is so high, the premium for the rapid response pales in comparison.

One example is a high precision machine shop that was asked to provide a part for a multi-million dollar aircraft in one quarter of the typical lead time. The customer was happy to pay \$800 for the part that was regularly \$200 because it meant they could ship the aircraft to their customer in 2 weeks instead of 8 weeks. Note that there is no significant change in the manufacturer's operating expense, typically. The manufacturer is doing exactly the same work, for a lot more money.

The implication of rapid response to the manufacturer's bottom line is momentous. In [Table 18-3](#), the example shows the difference between having regular sales at regular price, and having a rapid response offer that is taken up on 10 percent of the sales value, at a 50 percent premium of the regular price (on average). While total sales only grew 5 percent or \$50,000 increase on \$1,000,000 (\$100,000 worth of regular sales were billed at a premium of \$150,000), Throughput grew by 10 percent and net profit grew by 50 percent.

	Regular	10 % Rapid Response Sales at 50% Premium
Sales	\$1,000,000	\$1,050,000
Throughput	500,000	550,000
Operating Expense	400,000	400,000
Net Profit	\$100,000	\$150,000

TABLE 18-3 Effect of Rapid Response on Bottom Line

From experience, the greatest challenge with rapid response offers to the market is not in operations. With the processes already established in implementing the left-hand side of the S&T, my experience is that operations needs little time to do the additional improvements required to accelerate production lead time. The real challenge is in sales and sometimes in management, where the common belief is that no customer would ever pay more for the same product or service. Some inside the company believe that such charges are gouging a customer. Therefore, it's important to understand why this is not gouging and how to present it to the customer in a way that the customer perceives it positively, not negatively.

The correct strategy to roll out a rapid response program with positive customer perception is to test it in advance, qualify customer needs in terms of the potential quantifiable benefits of rapid response, and notify customers in advance of their potential needs. The customer must feel as if they have a choice and are not being blackmailed into paying more than necessary when they get into a bind. If a long-term customer of a manufacturer calls at the last minute and says, "Our forecast was wrong and I need another 100 parts in a hurry," and the manufacturer, without pre-notification, says, "I'll be happy to provide those parts in a quarter of the lead time for four times the price," they should expect a violently negative customer response.

If the customer is given the choice at the time original prices are quoted, customers are typically receptive to paying a premium if they have an emergency and they need the product. I would not believe this myself unless I had personally witnessed its success with a variety of companies. We have now taken several pages to discuss partially just a few of the slides in the S&T. This should convey a perspective of the depth of thinking contained within the tree. For further details, download the free S&T viewer and review the slides in detail, level by level, at least

10 times. From experience, each time you review, you will recognize new statements that you will swear were not in the S&T tree the last time!⁸

Projects

Once again, the S&T viewer described previously provides a full view of the entire project S&T and the TOC project paradigm called *Critical Chain Project Management* (CCPM). This S&T tree is intended for project companies (companies that earn the bulk of their revenues from executing projects for customers). In addition, every organization can use the majority of S&T outlined in 3.1.1 (view with S&T viewer) and below to improve their ability to execute change through projects more quickly.

The underlying TOC strategic thinking about projects is that almost all projects are conceived to bring some benefit to an organization.⁹ Every organization must execute multiple projects every year to survive and, we hope, improve. In most cases, the faster the projects are completed, the sooner the organization realizes the benefits. The faster the execution (flow of a project), the sooner the resources can be released and therefore the greater the number of projects an organization can complete with the same resources.

As we begin a journey to improve strategically how we execute projects, we first do Step 1 of the 5FS—Identify the System's Constraint. As hinted at previously, the biggest leverage point must be related to increasing project flow—the velocity at which projects are completed. The physical representation of a project is in the tasks needed to complete it. However, not all tasks equally determine the duration of a project. In defining which tasks are most likely to determine duration, TOC defines the tasks that are dependent on each other to complete. TOC recognizes two

types of dependencies—*logical dependencies* (e.g., the carpet installer will not install the carpets until the drywall is up and the rooms are painted) and *resource dependencies* (the carpet installer will not install the downstairs carpets until they're finished installing the upstairs carpet—the same resource is doing both tasks).

To summarize, *the longest chain of dependent events in a project, considering both task and resource dependencies, is the Critical Chain*. There is a huge problem with this definition of the physical constraint of a project. Some of the resources that are needed by a project in order for the project to progress are not modeled in project plans. In fact, projects always seem to progress at a rate that is much slower than the “touch time” or addition of the estimated task times for the critical chain. This revelation leads to two possible strategic assumptions we can use to *exploit* and *subordinate*:

1. The constraint is really management/support group attention. That is, the biggest gaps in project execution times are reflected in non-modeled resources, whose time and attention are not visible in any project plan. In this case, the fewer the number of active projects (to a point), the more attention that management and support groups can give to the active projects. This will result in fewer gaps in project progress, and gaps will be of shorter durations, thus accelerating the project completion. The POOGI process focuses on a Pareto analysis of what project tasks are waiting for, and focuses on constantly improving the project planning and execution processes in a way that reduces project durations.
2. A different strategic assumption is that the biggest constraint is still caused by behaviors and capacity of one project resource group or skill set. Still, if we

reduce the number of active projects so that this resource group can concentrate on a smaller number of projects, their work will progress faster, they will receive better support from other groups, and everything completes faster. The potential danger in this direction for a solution is that the organization assumes that by elevating this resource pool by adding more resources, this will increase the capacity of the organization to do more projects. If the real constraint is caused by lacking management or support group attention, then increasing the number of active projects without increasing management or support capacity will actually decrease project flow of all projects.

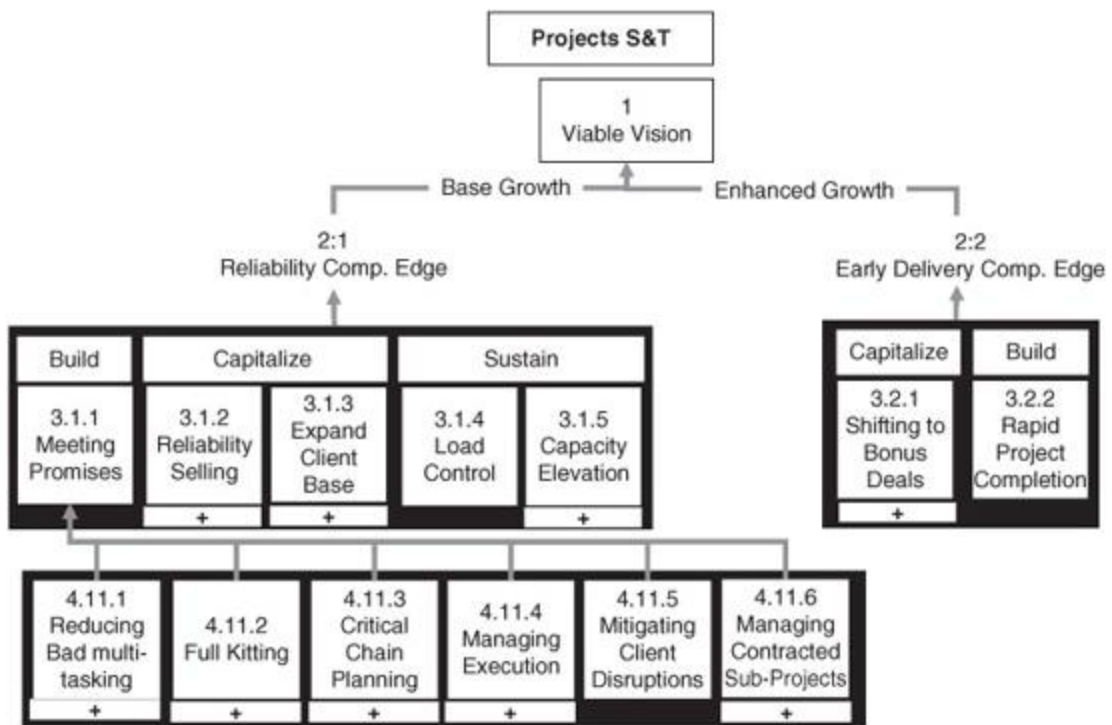


FIGURE 18.5 Projects S&T partial structure to four levels. (Used by permission of E. M. Goldratt 2008. *Projects S&T* © E. M. Goldratt.)

No matter which direction for a solution is chosen, the correct first step in deciding to exploit the constraint is to reduce the amount of active project *work in progress* (WIP). Therefore, you can see in [Fig. 18-5](#), in order to achieve “Meeting Promises” shown in box 3.1.1, the first step shown below this box, 4.11.1, is “Reducing Bad Multitasking.” This slide is shown in [Table 18-4](#).

Of particular note in slide 4.11.1 is the wording of the strategy—“Flow is the number one consideration.” Flow is measured by how many projects are completed each period. To achieve this, the tactic calls for controlling the number of active projects. Using the S&T viewer, you can view level 5 details, which call for freezing a substantial portion of the active project work. This is done before any critical chain plans are formulated or before any new software is used. This is how TOC seeks quick results.

To achieve results from the actions, the project S&T ties key elements together through a comprehensive cause-effect logic. For example, simply reducing active project work can easily result in no visible improvement. I know of an organization that cut their number of active projects in half, only to have the CEO remark that no one saw any effects. The reason is simple. All the resources were already so busy on the first 11 of their top 40 projects that no tasks on the other active projects were being worked on.

Therefore, if you view level 4 of the S&T (see [Table 18-4](#)) through the viewer referenced previously, you will see that immediately following the freezing of project work, the freed up resources are concentrated on remaining open projects, in order to speed them up. If, because of freezing projects, no resources are freed up, then the benefit of freezing is moot. This is an example of why the S&T between levels are inseparable. When an implementation team does not understand the full cause and effect of the entire process, the team and the strategy often fail.

4:11.1	Reducing Bad Multitasking and WIP
Necessary assumptions	<ul style="list-style-type: none"> • When too many projects are executed simultaneously, many resources will find themselves under pressure to work on more than one task—bad multitasking is unavoidable. • Prolific bad multitasking significantly prolongs each project's lead time.
Strategy	Flow is the number one consideration (the target is not how many projects the company succeeds to start working on, rather it's how many projects are completed).
Parallel assumptions	<ul style="list-style-type: none"> • The statement, "the earlier we start each project, the earlier each project will be finished," is not correct for multi-project environments. (Not only the first elephant but also the last elephant will go through a door much faster if they go in procession.) • Vast experience shows that in multi-project environments, reducing the number of open projects by at least 25 percent reduces bad multitasking without causing starvation of work and therefore significantly reduces the lead time of all projects—it increases the flow. (When most projects are of similar content, the rate of projects completed increases by a minimum of 20 percent.)
Tactic	The company properly controls the number of projects that are open at any given point in time. (The number of open projects puts less than 75 percent of the existing load.)
Sufficiency assumption	Adjusting the amount of work is not enough. The company must also ensure that as time passes the proper amount of work will always be maintained.

(Used by permission of E. M. Goldratt 2008. Projects S&T © E. M. Goldratt.)

TABLE 18-4 Projects S&T for 4.11.1 Reducing WIP

One of the most significant understandings drawn from the past few years of CCPM experience is that the key to driving benefits is not in planning projects, but in execution. Critical chain has a unique way of taking safety out of embedded task time estimates and accumulating the safety into buffers (project and feeding). This approach results in a shorter project timeline. However, at the time that tasks are estimated, there is a lot of vital task and project information that is still not known.

Three strategic execution processes drive down task times during execution. Without execution processes, even with full top management support, it's likely that a critical chain strategy will produce disappointing results. All of these processes are highlighted in the project S&T. The processes that are essential to achieve the strategy are as follows:

1. Full Kit—A great deal of project rework occurs because detailed work is started too early—before overall design/definition is complete. Generic full kit points and templates are defined for project types, which are then custom defined for each project. Combine this with the appointment of one of the top executives as a “Full Kit Manager,” and no follow-on work will be allowed to start until the full kit is reached. Unnecessary rework is driven down significantly. In a matter of a few weeks, resource time is freed up and costs (overtime, subcontracting, errors, etc.) are reduced.
2. Daily Task Management—The common practice in project management today is infrequent (weekly or longer interval) reviews of task progress. That is a long time to wait to find out that a task execution has a problem or is waiting for something to progress. Task managers, usually from within a skill set discipline, are given the full-time role of monitoring active project tasks daily. They prepare tasks in advance, ensure issues are resolved, and lend expert assistance to those executing the tasks. Their intuition and experience prevents over-engineering. Their practice of *not* assigning new work to a resource until a task is completely finished prevents multitasking—one of the biggest wastes of resources.

3. Fast Track Issue Resolution—A top manager is given responsibility to get all project issues resolved within 24 hours that could not be resolved below their level, either by project or task managers. This involvement of top management speeds up projects and brings the right level of executive attention to critical issues delaying projects.

Similar to the manufacturing rapid response S&T, the project S&T includes sales and marketing components to take advantage of the company's ability to execute projects more quickly and more reliably. A project company that can execute projects 25 percent faster than a few months ago must have a corresponding marketing and sales effort, in order to capitalize on this increased capacity.

The S&T also suggests that some clients are willing to pay premiums for faster project execution, even though the project company may not be willing to guarantee faster delivery. For example, if the project involves maintenance and repair of a very expensive asset (an aircraft or ship), the faster it's turned around, the less the revenue loss and the fewer the number of aircraft required in a fleet. Once the company has fully implemented CCPM, including its embedded POOGI, it becomes easier to beat the competition consistently on aggressive promise dates without taking risks.

In summary, CCPM processes in executing projects bring a level of cross-functional alignment, speed, and predictability that is very difficult for competitors to copy. Extensive top management support and discipline is essential to make this work. Once implemented, an organization capitalizes on the improved project execution capacity by value-oriented marketing and sales efforts. The impact on the bottom line becomes more and more significant, since some of the same project work can be done at a premium when it's executed much more quickly. Reading through the entire set

of slides in the S&T several times will bring the full picture to light.

Distribution/Retail

The average life of a golf club is 9 months. In bowling, specialty balls have a life of less than 6 months. Many medical products, including surgical rubber gloves, have a life of less than 2 years. We are not talking about shelf life or the product physically decaying. What we are describing is the perception created by the entry of new or replacement products, making existing products seem obsolete or not as valuable. No wonder distributors and retailers hate carrying large amounts of inventory! At the same time, retailers and distributors have the opposite paranoia—not having enough inventory and losing a customer sale.

Handled correctly, shorter product life is an excellent opportunity to enhance everyone's sales and profits. However, attempts at optimizing cost encourage mountains of unneeded inventory in the supply chain while simultaneously perpetuating shortages.

The TOC strategy for distribution and retail increases Throughput of the entire supply chain in at least six different ways (described in the next section). Often, the potential impact on business volume and profits is greater from Throughput increases than from reductions in Inventory. Thank goodness, the distribution solution does not force a choice—increased turns, increased Throughput, and decreased Inventory occur from the same holistic approach.

The S&T free viewer, described previously, should now be used to view the consumer goods S&T.

The Current Supply Chain Frame of Reference

To understand what determines a good strategy for distribution and retail of consumer goods, it's critical first to understand traditional practices. Driven by local optima, manufacturers push inventory into the distribution channel as soon as it's produced. Cost accounting rewards them in the short term by reporting the inventory movement as sales and profits on their books, even though no consumer has bought the products.

The Inventory has moved from the manufacturer's docks to the distributor's warehouse. The longer the Inventory sits in the distributor's warehouse, the higher the risk of obsolescence, the greater the carrying costs for the distributor, and therefore the lower that distributor's profits. For the distributor, profit is directly tied to the number of times they can turn the Inventory. Therefore, the manufacturer, the distributor, and each subsequent link keep pushing the inventory until it reaches the retail shops (if this link exists) or the end consumer level.

In this paradigm, much or most of the Inventory sits at the retail level, while little or no Inventory is at the manufacturer's plants. At the same time, customers who visit a retail shop often cannot find the specific item for which they're looking.

Why do some retailers carry so much Inventory if it just sits in their store much of the time? Why do they tolerate costly obsolescence and high carrying costs and still have customers walk out of their shops without the desired goods in hand? To answer this question, consider common manufacturer and distributor practices.

Many manufacturers and distributors offer quantity discounts, on a per-order basis. In addition, many of these same manufacturers and distributors have freight policies and charges that penalize the retailer for smaller orders. Since most retailers compete with other shops close by selling similar goods, the retailer needs a similar cost basis to stay competitive in price and margins. Therefore, retailers

place large orders with manufacturers (for much more than is needed to satisfy short-term demand) in order to gain better per-order discounts. As a result, retailers have much larger quantities of goods on hand than they need to cover immediate consumer demand and replenishment time. With this inventory on hand, even as consumer tastes are changing or as manufacturers replace existing products with new products, retailers are pushing *their existing* inventory on the consumer. You can see this phenomenon in the car market, computers, cell phones, etc.

Knowing the new product is coming, the retailers rush to get rid of the old products at fire sale prices or with special deals, anxious to avoid obsolescence. By doing these mass promotions, the retailers kill the market for the new product and their profit margins. Many of the consumers who might have bought the new product are now in possession of the older one.

There is another devastating effect to a supply chain that operates this way. By pushing most of their inventory to the retailers, the manufacturers and distributors have distanced themselves from changing trends in their markets. In many cases, with several months of inventory tied up at the retail and wholesale level, it takes manufacturers and distributors several months to see consumer demand changes reflected in orders to them. The lower the number of turns at each level in the supply chain, the longer it takes for manufacturers to understand and react to trends.

There is one other very negative effect. At the same time that excess inventory abounds in the supply chain, customers come to buy a product and often don't find it. Yet many times, the very product that is short in one store is sitting in abundance on a shelf somewhere else in the supply chain! This happens because of the practice of pushing large quantities of each product toward the retailers *in a manner that does not match end consumer demand*.

In this world of local optima, manufacturers keep on producing and shipping goods to distributors, sometimes “threatening” the distributor with higher prices or loss of exclusivity if the distributor will not accept the goods. Distributors do the same with their retailers, until the supply chain is completely clogged with goods that consumers are no longer buying. By this time, significant damage has been done. There is too much inventory in the system, and the part of the supply chain that has the most inventory suffers the greatest consequences. Retailers go bankrupt. Distributors lose money. Manufacturers—well, just look at what has happened in the great recession of 2008–2009. The auto industry is just one example of a whole supply chain exploding with too much inventory of what consumers don’t want.

Distribution S&T

As discussed in the previous example of the 5FS, the biggest leverage point is typically “the customers who come to buy the product or service.” There are not enough clients who buy.

We see distribution channels going adrift in Step 2, Exploit. Before spending money to attract more clients, don’t waste the ones who already come to buy. It’s a complete waste when clients don’t buy because they cannot find a product in a given location, but the same product is on the shelf in another location. In order to avoid wasting a customer, we must dramatically increase the chance of matching the right inventory in the right place, at the right time to the end consumer demand. To achieve this, we must answer the following questions:

1. Where is the ideal place to have most of the inventory?

2. What are the correct logistics to replenish the inventory between manufacturers and distributors, and between distributors and retailers?

We will begin with the first question—where the inventory should be located. Consumer demand varies widely from one geographic location to another. In any one retail location, it shifts dramatically from day to day. There are time lags between changes in end consumer desires and the reactions of the supply chain. All of these factors make predicting demand for any product at the retail level a phenomenal challenge.

Manufacturers often try to react to the challenge by pouring millions of dollars into more sophisticated forecasting systems, only to find little, if anything, improved. After all, a forecasting system does not make the end consumer react more rationally or predictably. To exploit the constraint—the customer who wants to buy—we must move away from sophistication into a much simpler solution.

There are two steps necessary to solve service level, inventory, and obsolescence night mares permanently in a distribution environment:

1. Pool the inventory where there is the greatest predictability.
2. Implement a pull system to replenish individual items frequently based on what was sold during a short interval.

If you were the logistics manager of a large shoe manufacturer, what is the ideal way to distribute the inventory of size 8 brown shoes across the supply chain? In one week, you could not say that anyone will buy even one pair of those specific shoes in any given store. However, you

would have much more confidence about the national numbers. *The more macro the level you are forecasting, the greater the predictability.* Predictability decreases as you move from national to regional and from regional to city and city to individual location.

Based on the characteristic of predictability listed previously, the logical thing to do would be to have more of the inventory at the manufacturing plant, less at the distributors, and even less at the retail level (see [Fig. 18-6](#)). Logically, you would also replenish with much shorter lead time. In this way, there will be less chance of stockouts at any given location and trends will show up much more quickly, with less waste.

For example, assume that a retail shoe store today is currently holding 3 months of inventory of each SKU. This inventory requirement is calculated based on the total replenishment lead time. The *replenishment lead time* is made up of:

- *Transportation lead time*
- *Production lead time—the time it takes the manufacturer to produce the product*
- *Order lead time—the time between when the retailer sold the first item and the time they reorder*

While the supplier controls transportation and production lead time, the retailer controls order lead time. As it turns out, order lead time usually provides the biggest opportunity for improvement. In the old approach, retailers reorder when they hit a minimum quantity in stock (commonly called the min in a min/max system or reorder point in a reorder point/economic order quantity system because they will order sufficient quantity to bring their

stock back to a maximum level). Under the exploit step, TOC implements a pull system that has the retailer order, every period (e.g., one week), exactly what they sold in the previous period.¹⁰

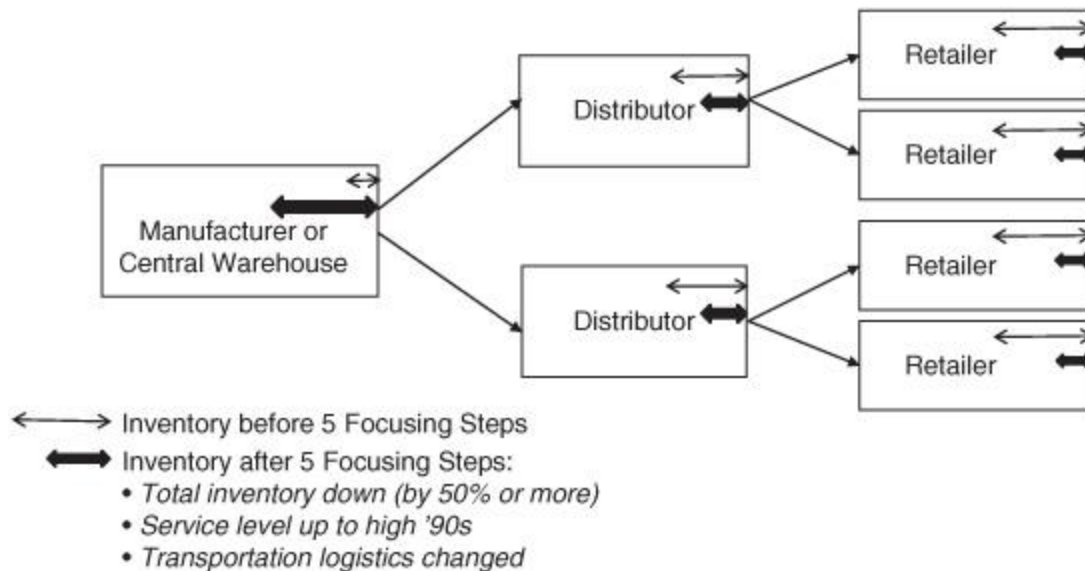


FIGURE 18-6 Inventory impact of TOC Distribution Solution. (Used by permission of E. M. Goldratt 2009. *Distribution S&T* © E. M. Goldratt.)

In our new system, assume transportation lead time stays the same. With finished goods inventory now being held by the manufacturer at the plant warehouse, production lead time is 0—the product should be in stock in the plant warehouse. The biggest difference is usually in the order lead time. Instead of waiting 2 to 2.5 months for the stock level of each item to hit a minimum, and ordering a 2 to 3 month supply of that item, the new pull system has the retailer order only the SKUs that moved the previous week, in the exact quantities sold the previous week. The order lead time has now been reduced from months to a matter of days.

This means that the amount of inventory of each item that the retailer or distributor needs to carry to cover their

replenishment lead time is much less. It also means that their chance of running out of an unexpectedly popular item for weeks at a time is almost nil. Service levels, under this new system, move much higher, while total inventory in the supply chain typically drops by two-thirds! The system reacts much more quickly to the variability and uncertainty of consumer demand. It also greatly simplifies both the distributor's and the retailer's lives. Instead of worrying about ordering a 3-month supply and debating how much and which products the consumers might demand over that long a time period, the order items and quantities are automatically determined based on sales.

Wearing our marketing hat, with the additional space freed up in the store by reducing the overstocks, we would have the retailer carry a greater variety of our goods. More variety means more sales (see [Fig. 18-7](#)). The left side of the S&T focuses on TOC replenishment throughout the supply chain, from manufacturing through distribution and retail. The right side of the tree focuses on Throughput per square foot or meter of shelf space (abbreviated TPS—*Throughput per shelf*).

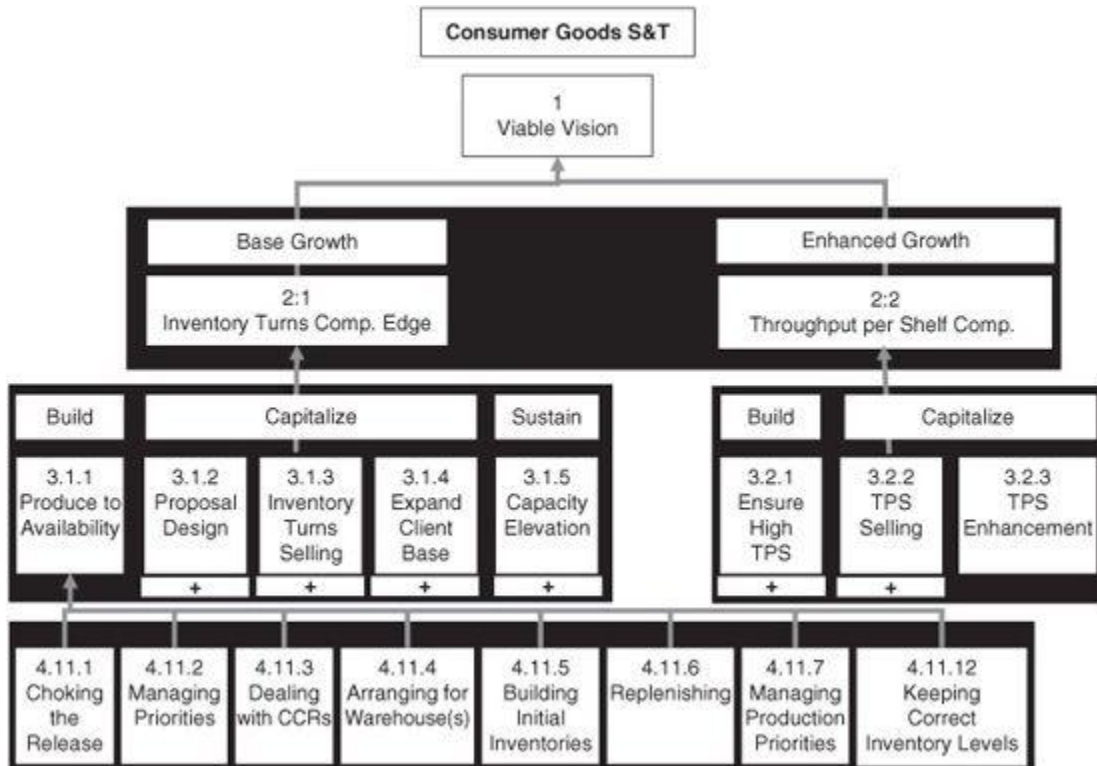


FIGURE 18-7 S&T for TOC Distribution Solution. (Used by permission of E. M. Goldratt 2009. *Distribution S&T* © E. M. Goldratt.)

As for transportation costs, in the real-life cases where this logistics solution was implemented, the configuration of each of the shipments was different, but the frequency and cost of shipping remained about the same. Instead of shipping 50 units (a 3-month supply) of 10 items, we are shipping 10 units (a 3-week supply) of 50 items. Of course, shipping more frequently in some industries may well be worth an additional cost (if there is any) with a disproportionate increase in sales and decrease in obsolescence.

For distribution chains that have changed logistics according to these principles, the statistics are astounding. Overall OEs (considering the combination of transportation costs, inter-warehouse transfers, inventory carrying costs,

returns, and obsolescence costs) decrease dramatically while Throughput increases significantly.

Successful companies don't stop at the exploit step. By focusing attention on Step 3, *subordinate*, you substantially reduce any risk of failure.

Subordinate means that everyone in every part of the distribution supply chain, including the manufacturers, adopts the attitude that "As long as the end consumer has not bought, we have not sold!" Existing measurements must change to encourage pull rather than push. This is especially important for the manufacturer, who will now be carrying a significant portion of the total finished goods for the entire supply chain in the plant warehouse. See the S&T for consumer goods for a full insight into this approach and for the assumptions made at all levels of the supply chain. [Figure 18-7](#) provides an overview of the structure.

In the subordinate step, the company puts the pull system software and procedures in place. The minimum inventory necessary to cover fluctuations in consumer demand and transportation time during a short time period are stocked at the retail level (or level closest to the end consumer). The distributor carries the minimum inventory necessary to cover fluctuations in demand from their customers plus the transportation time from the manufacturer to replenish the inventory. The manufacturer carries a larger inventory to cover fluctuations in distributor demand plus the cycle time to manufacture sufficient quantity to replenish their plant finished goods inventory. Of course, you must factor these numbers based on the reliability of the transportation and the manufacturer into these calculations.

Some of the most successful companies in the world are using such systems today—from computer manufacturing to super retail stores to luggage companies. Just look at the inventory turns of highly profitable manufacturers and distributors, and the numbers speak for themselves. They have implemented a true pull system.

What about Step 4, *elevate*? Distributors might need to elevate, for example, if they want to open up new geographic areas. However, with the cash freed up from the first three focusing steps, it's much easier to add new distribution locations. In addition, the cost of setting up each new distribution location is much smaller, due to the decreased inventory.

Six Ways That the Holistic Distribution System Increases Throughput

1. Having sufficient stock of each product at the right location in time to match demand increases the service level (the percentage of time the buyer finds the item in stock and therefore is able to complete a purchase).
2. With less inventory, and therefore less shelf space required per item, the distributor and retailer are able to stock a greater variety of items, thus increasing Throughput from the same physical space.
3. With less inventory per item, sales to get rid of excess inventory are required much less frequently. This means that both the margins and the sales revenue are increased. It also implies that sales of older products don't spoil the market for new products.
4. With less inventory per item, there is less obsolescence, leading to greater satisfaction among consumers and therefore increased sales. The products on the shelf are newer and therefore more appealing. This can have a huge impact on Throughput with items that have expiration dates.

5. Faster replenishment during out-of-stock situations means that there will be much less opportunity for a buyer to have to resort to a competitor's product.
6. The faster reaction to changes in consumer demand leads to fewer and shorter times of being out of stock of any given item.

In summary, many distributors fight hard with their manufacturers to carry smaller quantities of inventory on better terms. At the same time, the manufacturers are often trying to do the opposite—push larger amounts of inventory to distributors and retailers with more aggressive terms. Each link tries to push their inventory as quickly as possible to the next link in the supply chain. This creates huge gluts of inventory at the retail level. As retailers try to predict, months in advance, what their consumers will demand and order accordingly, it's 100 percent certain that retailers will have too much of some products and not enough of others during this long time period.

TOC changes this nonsense from a push system to a pull system. Inventory is kept where it makes sense (less at the retail level,¹¹ more at the manufacturer's and regional warehouses). This is entirely feasible when the time to replenish goods at the retail level is significantly decreased. With these changes, total inventory in the supply chain decreases by two-thirds, while customer service increases.

Four Generic Prerequisites/“Injections” for a Lasting Competitive Edge

There are four prerequisite conditions to achieving any good strategy quickly:

1. A common, correct frame of reference for all management, including a global measurement system that induces and encourages holistic behavior across the supply chain. The TOC framework is driven by holistic measurements (Throughput, Investment, and Operating Expenses) combined with the 5FS.
2. Stable and predictable operations logistics that allow the organization to “turn on a dime.” An organization must be able to meet their customer due date and lead time expectations. Further, as demand increases by large amounts, the organization must be able to respond quickly. If a company responds to a large increase in demand by temporarily extending lead times by a short factor, most customers will find this acceptable. By contrast, if a company extends lead times by more than 15 percent for more than a few weeks, it’s predictable that many customers will start to look for another source, unless the phenomenon is industry-wide. To prevent extending lead times by a large percentage for a long time requires a focus on one or two variables—the leverage points in the operation that must be adjusted quickly. For example, within a forging operation consisting of dozens of production departments, the two leverage points might be in the heat treat ovens and at the forging press. In a labor-intensive operation such as building custom kitchen cabinets, the two leverage points may be in the spray booths, where it’s difficult to train and retain labor, and in the hiring process. A logistics solution that includes quick capacity elevation preparation *must* be in place and *in alignment with* the valve to the market.
3. Correct and sufficient marketing. There must be recognition that scarce marketing resources must

consciously focus on the few target markets where the company can excel. The implication of scarce sales resources is that lead generation and sales cycle management is a must in a value sale.

4. The ability to implement change quickly and predictably. This means superior project management execution. TOC's Critical Chain Multi-Project methodology provides the means to drastically cut project durations and increase predictability of successful completion. Critical Chain is not optional—it's a prerequisite to achieve a VV.

Once an organization has implemented these conditions, it's stable and can adapt very quickly to market changes. Here are four strategies (injections in a TOC FRT) that, in combination, are designed to lead to market domination for 10 to 15 years:

INJ. 1: Increase Customer Perception of Value that Competitors Have Difficulty Copying

The company implements modifications to current product/service offerings that substantially increase the customer perception of value for a sizeable market. This injection uses the concepts of the Mafia Offer. To create such an offer, you must really understand the customers and how they will benefit significantly from changes in:

- Options
- Packaging
- Service level
- Guarantees
- Response or lead times

- Removal of industry standard annoyances, such as freight charges, minimum order quantities, etc.

Don't make such an offer based on lower price. Price is the easiest attribute for any competitor to quickly copy. You should look at rooted industry policies first, before changing the physical product itself.

The result of implementing this injection is a significant advantage in a sizeable market. In addition, the marketing department must create marketing awareness/presentations that reposition their products with higher value to the market. The salespeople must learn how to sell value effectively. The operations people must have processes in place to manage significant increases in demand. These successes help to build the confidence of all of these functional areas in the strategy, a desirable condition before implementing injection 2.

INJ. 2: Implement Practical Segmentation

A market is considered segmented when the prices and quantities sold in that market segment have no impact on the prices and quantities sold in any other segment. Such segmentation gives you the opportunity to address the needs of different groups of customers in a unique way, with the same base product or service. For example, a manufacturer of expensive shoes enters the mass-market segment for less expensive shoes. A maker of original truck parts segments by entering the aftermarket for replacement parts. A distributor of advanced products in the wireless communications industry uses their knowledge to distribute basic consumer goods.

The result of the implementation of this injection is that the company is operating in several market segments in

which it has a competitive edge. The company elects not to take 100 percent of any market segment that is not very lucrative. A company that owns 100 percent of a market segment has much less flexibility to improve. Being in a monopoly situation implies certain responsibilities. If you decide to get out of that segment, and you don't leave your customers with an alternative, those customers will hate your organization for a long time. Don't put a lot of energy into taking 100 percent of a market unless it's hugely beneficial. Conserve your resources for where they will do the most good.

The company is careful to enter only into new products that require almost the same resources (people) as it already has. A company that can shift its resources at will, easily, between different markets and opportunities is a company that has excellent flexibility. Such flexibility gives the company the ability to meet the necessary condition of employment security and satisfaction in a way that makes more money for the company. The key point to remember when implementing this strategic injection is that it's *people*, not machines, about which we are talking. A person who is a great manager of engineering airplane components, for example, also can manage many other engineering environments.

Because of implementing this injection, the company succeeds in segmenting its markets, not its resources, with flexibility to shift resources at will.

In choosing its overall collection of market segments, the company looks for segments where the probability of many segments experiencing economic downturns simultaneously is very small. When combined with injections 1, 3, and 4, this one provides the company with the insulation it needs for long-term stability. When a lucrative segment is up, by deliberately *not* taking 100 percent of any other market segment, the company has the flexibility to shift resources from less desirable segments. When a segment is down, the

company can shift resources to other segments. Given this situation, the company rarely will find itself in a situation where it's forced to lay off people.

INJ. 3: Identify and Build the Decisive Competitive Edge Factor

The company uses the window of opportunity created by implementing the previous injections to identify a factor in which an order of magnitude improvement will bring a significant competitive edge, and uses its improvement efforts to achieve it. With injections 1 and 2 successfully implemented, it's possible that competitors will catch up within 2 to 3 years. This injection calls for the company to use this window to build a decisive competitive edge. The company now must identify *one* factor and develop it to be several times better than current performance. For example, in computers, this would be a machine that operates at 5 times the current performance or 5 times the simplicity. In air travel, imagine an airline that could get you to your destination in a quarter of the time it now takes you. It does not necessarily mean that the airplane is 4 times faster. In building a car to order, instead of 12 to 15 weeks lead time, picture such a custom-built car delivered to your door within 2 weeks. Such factors exist in every industry, but require turning a company on its head to make it a reality. It's not only a technological challenge, but also it usually means aligning efforts in engineering, production, distribution, and marketing.

Because of this step, the company now has the potential to increase revenue well beyond its current capacity.

INJ. 4: Strategic Segmentation

Within most target markets, there is a subset that values the factor you have developed much more highly than other

segments. Strategic segmentation means that you can price the identical product at a premium, when adding a guarantee or service component based on the factor identified previously. Segmenting the markets and pricing services according to this customer perception of value is the key to achieving “impossibly” high profit goals. For example, take an 8 oz. envelope and ship it with guaranteed first overnight delivery, and the fee is \$44.10. Take the same envelope, using the same airplanes and trucks, but choose the standard overnight service, and the fee drops to \$16.55. Put the same envelope in the mail, and the fee is 44 cents. Similarly, we have proven that some customers will pay 50 percent more for the identical custom kitchen when it’s delivered in 2 weeks instead of 6 weeks. A manufacturer of commercial aircraft will pay 4 times the standard price for delivery of critical items (e.g., custom painted emergency door handles) when it will help that manufacturer ship the aircraft a few weeks earlier.

From my observation, most companies have such market segments but fail to distinguish them in their service offerings and pricing. As a result, they have some customers who perceive their product to be “expensive” or too high priced, relative to their needs, while at the same time other customers consider the product to be a fantastic bargain.

How do you go about identifying such segments? The answer is that you must get to know your customer’s business better than the customer themselves know it. For example, in the case of the custom kitchen manufacturer, the VP of sales and I visited several dealers, and asked them if they ever got requests for expedited deliveries. “Of course,” was their answer. Sometimes, the customers (often contractors who had gotten themselves into trouble in the construction process or the dealer who took too long to finalize an order) simply tried to put pressure on the manufacturer to deliver earlier. The guilt trip of “You know how much business I give you every year?” often worked,

but to the manufacturer's disadvantage. By expediting one customer's order without the proper logistics in place, it often meant making another customer suffer or incurring extra costs of overtime and expedited freight.

The revelation was that when the proper logistics and improvement on production lead times was put in place, the manufacturer could deliver 20 percent of their orders in a quarter of industry standard lead time without extra costs, and some part of the market was willing to pay a handsome premium for the shorter lead time. There is no question that these segments exist. The only question is how big they are within given target markets.

The ultimate in exploiting this strategic segmentation is in a sales strategy where salespeople are dedicated to finding these premium deals. In these cases, the manufacturer can reserve capacity for premium segments, and double or triple their profitability by concentrating on finding those prospects in the market whose perception of value for the premium service is high.

Desirable Effects of a Good Strategy

Within every TOC strategy, there is an explicit set of desired effects (DE) defined. We must achieve these results. Ideally, DEs are measurable, tangible conditions. For purposes of strategy, we want these conditions to perpetuate indefinitely.

In the TOC FRT structure, these are explicitly listed as DEs. Within an S&T structure, these are stated as strategies (what we are striving to achieve) at each level. For example, if you look at the manufacturing S&T, level 4 under the 99 percent DDP, there are three key strategies that you can consider as DEs:

1. The shop floor is populated only with orders that need to be filled within a predefined horizon.
2. The shop floor is governed by a simple yet robust priority system. For example, in one company with several hundred large pieces in WIP at any point in time, red, yellow, and green tags were placed on all pieces. The shop floor operators could choose any one of the pieces of the highest priority color to work on. When following the first principle of limiting WIP, the operator typically had a choice of only one or two pieces. The operators and their supervisors claimed that this simple visual system was the best one under which they had ever worked. Every shift, colors were changed on the few pieces that had moved from one status to another.
3. Orders are shipped on time (>99 percent, for example).

If we achieve these three DEs, and have these conditions existing on an ongoing basis, then we should have the DE shown as strategy in the next higher level 3.1—the company has very high DDP (> 99 percent, for example).

To achieve the three necessary desirable effects of a good strategy—goal units are increasing now and in the future, employee security and satisfaction exist now and in the future, and the organization is satisfying its markets, now and in the future—the following generic DEs are the stepping stones:

1. Every major change effort achieves quick, measurable results (quick implies within 8 to 12 weeks).

2. The company has a decisive competitive edge within many (>5, preferably >10) market segments.
3. The company's employees are easily shifted between market segments.
4. The market is the constraint. This statement requires some clarification. The assumption has been stated frequently within TOC conferences that there is no real market limit to company growth. The world economy (with exceptions of brief periods over the past 200 years) continues to grow. With billions of new consumers just starting to enter the market with real buying power (e.g., China and India), the world demand for goods and services will grow exponentially. Therefore, market potential is not the constraint. When we declare the market constraint to be a DE, it means that we choose not to be constrained internally. We choose to expand our organization, at our will, based on the rate of growth that we believe is good for our organization.
5. The company has no monopoly in any product or service. This gives the organization the ability to withdraw or decrease service from a market segment without doing damage to their reputation. The company does not want a situation where they have customers dependent on them, and then they drop a product where the customers have no ready alternative.
6. Layoffs are rare. Never is preferable, but if a layoff is absolutely necessary due to cash flow threat, it's not repeated within a 5-year period.

Two Forms of Strategy and Tactics—TP and S&T Trees

Several books describe TOC TP and how to construct *Evaporating Clouds* (ECs, sometimes called conflict diagrams) and an FRT using TOC TP (see Dettmer, 2007; Scheinkopf, 1999; Chapters 23, 24, and 25 of this Handbook). The following brief discussion assumes that you already have knowledge of this subject matter.

The two different formats of TOC strategy (FRT and S&T) have been discussed and illustrated previously. In order to construct an FRT, it's usually necessary also to construct ECs in order to better understand and overcome the root cause of major system problems. In addition, ECs provide assumptions and injections that can help lead to a direction for a solution.

Mapping:

- S&T Tactic = Injection in FRT
 - S&T Strategy = DE in FRT
 - S&T Assumptions may equate to some entities in an FRT that build sufficiency in the cause-effect logic of an FRT
- S&T mapping to EC: An EC is a powerful tool, which many TOC practitioners use to better understand the problems and find directions toward a good solution. Several elements of an S&T can be discovered using such a tool. For example, ECs can be used to choose a strategy in an S&T (for example, an EC about different directions for a solution may point to one strategy over another). An EC can also help identify a tactic in an S&T (for example, in a conflict related to achieving the strategy, where the strategy is the common objective of an EC, one or more of the assumptions in the diagram may lead to a tactic to overcome the assumption). S&T assumptions may also be identified directly from assumptions in the EC.

There are different perspectives on the use and usefulness of the two different formats. For example, some people believe that the logic of a strategy is best developed using the 5FS and TP, and best communicated to others using the S&T format. My personal experience is that a strategy can be developed using either tool, depending on how your mind is trained.

The free Harmony viewer introduces the S&T format and brief instructions on how to construct an S&T from the beginning.

Integrating Other Methodologies Such as Lean and Six Sigma

To sustain any organization, TOC provides a significant part of the answer. Lean, Six Sigma, and other methodologies and knowledge complete the solution. Processes are needed to provide:

- Flow of the product or service quickly and efficiently enough to profitably meet customer demands (TOC provides logistics for flow).
- Quality sufficient to meet customer needs (Six Sigma is a common methodology used to increase quality).
- Efficiency sufficient to competitively meet customer needs (Lean is the most popular methodology for removing waste from a system).

While each methodology claims to provide benefits in all three necessary conditions, the strength of each, in my opinion, is as highlighted previously. The common positioning within TOC is that the constraint should guide

where to apply Lean and Six Sigma efforts. Lean and Six Sigma literature is filled with similar sentiments, that is, that those methodologies are the main ones, and any others should be subservient. The assumptions behind the TOC guidance are:

1. When other methodologies are applied everywhere, the resources needed to address a constraint end up being tied up. These actions distract from exploiting and subordinating steps.
2. When other methodologies are applied everywhere, much of it is a waste of time because benefits accrue to the company only if the actions result in increasing goal units of the company.

After over 40 years of business experience and 15 years of TOC experience, my current assumptions are:

1. We cannot predict, in advance, exactly when and where Lean and Six Sigma skills may be necessary to advance a project or help the organization to achieve its goals. These skills take time to develop and generally are not very useful without a person or group having some practical experience.
2. When there is a major increase in company flow, new challenges in quality and waste often arise, which threaten flow. This frequently happens due to hiring more new people than the company had in the past or due to exceeding the capacity to cope with quality issues. For example, if a shop floor lead hand is used to personally deal with 10 problems a shift, and flow doubles without changing quality processes, now that same individual is trying to cope with 20 or more problems per shift. As flow increases, without

increasing machine capacity or work force, the ability to deal with some underlying problems can create a bottleneck. TOC suggests applying Lean or Quality techniques now to unblock the flow. However, if these skills don't exist within the company, the organization may be at the mercy of an outside consultant's schedule and expertise to address the issues or at the mercy of a training program.

Due to these issues, I believe that it's a good strategy for a company to build these skills proactively, as part of their long-term investment in their people, much as Toyota has. If this approach is integrated with a TOC strategy, the results are more predictable and sustainable. There need not be any inherent conflicts between TOC, Lean, and Six Sigma when the three methodologies are applied within an overall strategy if one follows a Throughput World focus versus the traditional Cost World focus.¹² There is a great deal of potential damage (infighting, methodology zealots, and stagnation, for example) that is predictable if the organization executives don't establish such an overall, integrated approach up front.

Dealing with Human Behavior in a Strategy

What about the human side of strategy? It was stated earlier that employee security and satisfaction are a necessary condition of having an organization built to last. While important aspects of security and satisfaction can be achieved by an organization that continues to grow and prosper, there is more to satisfaction in today's knowledge worker age. While TOC has some ways to address human behavior with a set of processes called "Management Skills" (see the TOC TP books mentioned previously), there are

some other necessary conditions of executing a great strategy that remain unaddressed.

In a book called *The Speed of Trust*, Stephen M. R. Covey (2006) describes documented cases of the tangible cost of poor people practices. The speed of making changes and executing decisions is greatly increased in organizations that have high trust, a measurable parameter. Another group of authors (Patterson et al., 2002; 2004; 2007) wrote a series of books that describe scientific research and confirmation of how to influence human behavior and the cost of poor communications.

My experience is TOC strategies can be implemented at least twice as quickly with double the success rate when the organization has excellent communications to begin with. Many organizations suffer from communications issues, especially during a transformation process or periods of high growth. I believe it's vital for an organization to include the development of these skills as part of any strategy. My suggestions for accomplishing this part of the strategy are:

1. Choose one of the science-based behavior programs (e.g., Covey, Influencer) based on the most important current company needs. Read the books to determine which program best suits the current organization needs.
2. Set a tangible, measurable goal for the behavior changes desired.
3. Pilot the program in a functional area or department where the biggest need exists. Measure the before and after parameters.
4. Assuming success in the previous step, roll out the program across the organization, in as short a time as possible, starting with the top management team.

(One excellent way to kill the effectiveness of such a program is to start at lower levels and have people become discouraged because the top management is not practicing the principles.)

Summary

The real strength of TOC lies in the thinking that forces an organization to explicitly identify and focus on its biggest leverage point—the constraint of achieving the organization’s goal. TOC provides a strategic tool, the 5FS, to identify the constraint, and S&T Trees to detail and communicate the detailed steps and expected results. The TOC TP provides a way to overcome problems if you get stuck at any one of the 5FS. Any strategy can be expressed using one of the two TOC formats—S&T Tree or FRT. While some elements of each format can be mapped to each other, the detailed content and organization are quite different.

Generic TOC strategic and tactical solutions exist for common industry problems. Such solutions in the public domain (see Introduction for Website reference) exist for manufacturing flow, for distribution of discrete products, and for projects. All such solutions provide three essential elements—the logistics to build a decisive competitive edge, how to capitalize on it through sales and marketing, and how to sustain it with processes that deal with capacity issues.

Other methodologies, such as Lean and Six Sigma, can and should be integrated with TOC to provide a comprehensive solution to any organization’s strategic needs (See Chapters 6 and 36). The top management team must decide how to integrate methodologies to focus on Throughput, or risk confusion and in-fighting over which methodology is “best.” To execute strategic changes

quickly, without top management constantly force-feeding, human behavior and communication skills are essential. Today, there are proven scientific approaches to positively improve human behaviors.

TOC strategy, by itself, is not the complete answer to an organization's needs. At the same time, any organization without a TOC strategy is definitely missing a great deal.

References

Bossidy, L. and Charan, R. 2002. *Execution: The Discipline of Getting Things Done*. New York: Crown Publishing Group, Random House.

Bruner, R. F. 2004. *Applied Mergers and Acquisitions*. Hoboken, NJ: John Wiley & Sons, Inc.

Collins, J. 2001. *Good to Great: Why Some Companies Make the Leap... And Others Don't*. New York: HarperCollins Publishers, Inc.

Covey, S. M. R. 2006. *The Speed of Trust*. New York: Free Press.

Dettmer, H. W. 2007. *The Logical Thinking Processes*. Milwaukee, WI: American Society for Quality.

Goldratt, E. M. 1990. *What is This Thing Called Theory of Constraints and How Should it be Implemented?* Croton-on-Hudson, NY: North River Press, Inc.

Goldratt, E. M. 2008. Projects Company S&T, Level 5, July at: <http://www.goldrattresearchlabs.com>

- Goldratt, E. M. 2008. Retailer S&T, Level 5, July at: <http://www.goldrattresearchlabs.com>
- Goldratt, E. M. 2009. Manufacturing Make-to-Order (MTO) Reliable Rapid Response S&T, Level 5, May at: <http://www.goldrattresearchlabs.com>
- Kotter, J. P. 1996. *Leading Change*. Boston, MA: Harvard Business School Press.
- McDonald, J., Coulthard, M., and de Lange, P. 2005. "Planning for a successful merger or acquisition," *Journal of Global Business and Technology* 1(2)(Fall):1-11.
- Patterson, K., Grenny, J., McMillan, R., and Switzler, A. 2002. *Crucial Conversations: Tool for Talking When Stakes are High*. New York: McGraw-Hill.
- Patterson, K., Grenny, J., McMillan, R., and Switzler, A. 2004. *Crucial Confrontations: Tools for Talking about Broken Promises, Violated Expectations and Bad Behavior*. New York: McGraw-Hill.
- Patterson, K., Grenny, J., Maxfield, D., and McMillan, R. 2007. *Influencer: The Power to Change Anything*. New York: McGraw-Hill.
- Scheinkopf, L. J. 1999. *Thinking for a Change*. Boca Raton, FL: St. Lucie Press.
- Schneier, C. E., Shaw, D. G., and Beatty, R.W. 1992. "Companies' attempts to improve performance while containing costs: Quick fix versus lasting change," *Human Resource Planning* 15(3):1-26.

Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. *The Theory of Constraints International Certification*
<http://www.tocico.org/?page=dictionary>

About the Author

Gerald I. Kendall, PMP, Senior Consultant, The Goldratt Institute, is a recognized world expert at strategic planning, Theory of Constraints (TOC), and project portfolio management, with extensive implementation experience. His clients span the globe, including engagements in Malaysia, Bangladesh, Australia, Europe, the United States, and Canada. Clients include SAP, Telstra, British American Tobacco, Raytheon, Alcan Aluminum, Rio Tinto, and Lockheed Martin, with a variety of small company clients such as machine shops and manufacturers.

Gerald began his career with IBM as a systems engineer and became an IT Director. After expanding into international sales and marketing, with global executive responsibility, he broadened his experience in strategic planning, supply chain, and operations. He has implemented TOC solutions for manufacturing, distribution, projects, marketing, sales, people management, and strategy and tactics. He currently leads several multi-million dollar transformation/high growth projects using TOC. Gerald is certified in all disciplines of TOC by the TOC International Certification Organization.

Gerald's latest book *Denistry with a Vision* (November 2009), guides doctors, dentists and other professions in using TOC, Lean and Six Sigma to improve their practice. (October 2004) *Viable Vision* explains to executives and managers how to get high leverage out of an organization's change efforts. His *Advanced Project Portfolio Management and the PMO* is the top selling book in the PMO and Project Portfolio Management space. His first book, *Securing the*

Future, continues to be in high demand in its 11th year of publication. Gerald also authored a chapter on Project Portfolio Management for the American Management Association *Handbook of Project Management*, 2nd ed. , and a chapter on Critical Chain in Dr. Harold Kerzner's book, *Project Management, A System's Approach*, 8th ed.

Gerald is a graduate and silver medal winner of McGill University. You may email to Gerryikendall@cs.com. Web site: www.goldratt.com.

CHAPTER 19

Strategy

H. William Dettmer

The Popular Conception of Strategy

Everybody talks about strategy...

- “What’s your strategy for finding a job?”
- “What’s our strategy for getting this project done on time?”
- “What strategy can I use to get out of debt?”
- “What’s our strategy for winning the next election?”
- “What’s your strategy for getting your spouse to agree to our golf trip to Las Vegas?”
- “What’s the strategy for turning around the slumping economy?”
- “What’s our strategy for winning the game next Sunday?”
- “What strategy should we use to introduce this new product to the market?”
- “What strategy can bring peace to the region?”
- “What’s your strategy for getting Nadine to go out on a date with you?”

From this list, it should be obvious that the word *strategy* is used in many different ways to connote a wide variety of meanings.

Strategy's origin is military, dating back as far as the Chinese general, Sun Tzu in the 5th century BC (Cleary, 1991). In modern times, its military aspect is most often associated with Clausewitz, Moltke, Liddell Hart, and, more recently, Boyd. Nearly all military definitions of strategy involve objectives, winning, application of resources, and execution of policy.

The commercial business community tends to see strategy almost exclusively in terms of Marketing or Finance. Michael Porter's (1985) famous "low-cost leader versus differentiation" concept was the basis of his landmark book, *Competitive Advantage*, the virtual bible of business schools for many years. However, such a narrow characterization ignores the applicability of strategy to other kinds of activities and organizations, such as government agencies and not-for-profit groups—systems that do little or no Marketing and Sales, or are not in business to generate a profit. Moreover, it fails to consider some of the personal, but no less valid, applications of the concept.

Copyright © 2010 by H. William Dettmer.

The underlying relationship is *not* between strategy and a particular type of organization; it's between strategy and *systems*. Understanding the distinction frees the imagination from artificially imposed constraints on how, and for whom, strategy might be constructively employed.

The System Concept

It is difficult for many people to think conceptually in terms of systems. It's easier for them to pigeonhole systems as "organizations," either formal or informal. Yet, as [Table 19-1](#) shows, the system concept goes well beyond organizations.

In its simplest incarnation, a *system is made up of inputs, a process of some kind, outputs, and the environment in*

which these components exist (see [Fig. 19-1](#)).

Any system interacts with other similar (or dissimilar) systems that coexist in the same environment, and with elements of the external environment itself. Some of these other systems might include suppliers, customers, regulatory bodies, special interest groups, competitors, societal groups, educational institutions, etc. The interactions among systems—or lack thereof—are related to the nature of the system's chosen functions and activities.

In view of the far-reaching nature of systems and their interactions with other systems and the environment, it would be myopic to consider the concept of strategy exclusively in terms of narrowly defined organizations or departments such as Marketing/Sales or military operations. Moreover, while strategy can certainly be developed and deployed without any prior knowledge of the *Theory of Constraints* (TOC), a thorough familiarity with TOC concepts and principles, in addition to systems thinking, enhances the quality of any strategy subsequently developed. More needs and opportunities are likely to become visible.

A Vertical Hierarchy

Besides the “horizontal” conception of strategy across different types of organizations—commercial, not-for-profit, government agency—there's a vertical perspective as well. This vertical aspect is related to system *levels*.

Human	Economic	Political
Personal	Commercial	Governments
Family	Economies	Administrative
Society	Local	Political parties
Cultural	State	Revolutionary movements
Educational	National	Information
Charitable	Transnational	Security (law enforcement, military)
Social	Information	
Knowledge	Technical	

Note: Biological and other “non-thinking” systems are excluded from consideration here. Our attention is confined to systems involving human cognition and decision-making capability.

TABLE 19-1 Types of Self-Aware Systems

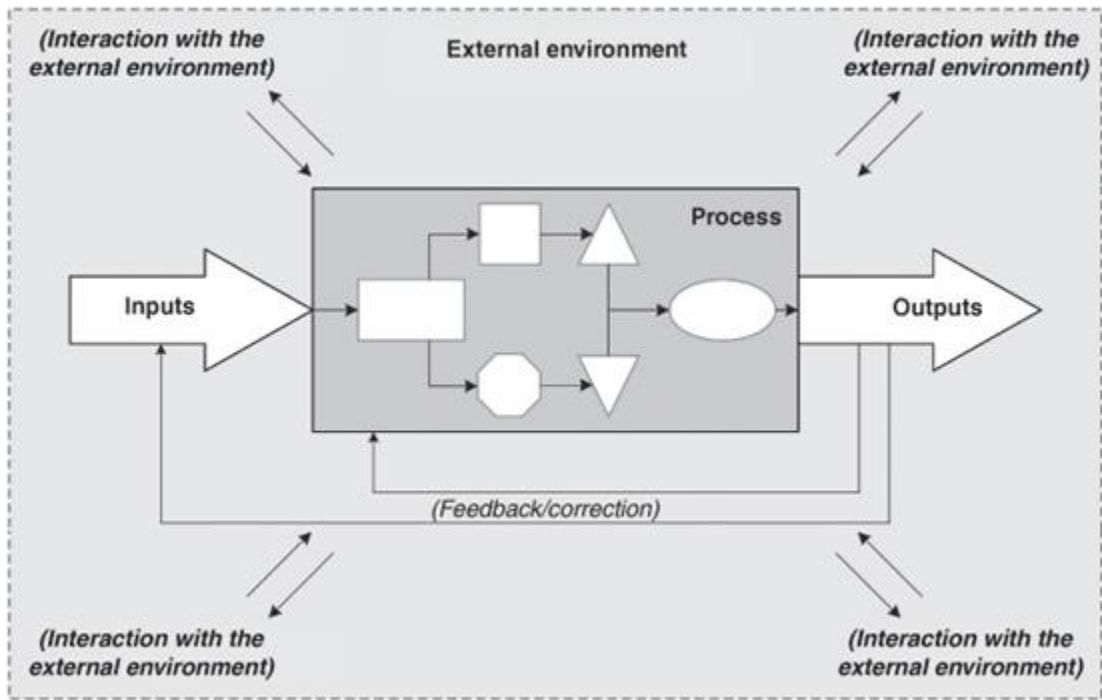


FIGURE 19-1 Basic system.

Systems are hierarchical. What usually occupies our attention is no more than one level of a larger system

composed of multiple levels. An old rhyme characterizes the vertical relationship:

Big fleas have little fleas
Upon their backs to bite 'em.
Little fleas have lesser fleas,
And so on, ad infinitum. (Ramel)

Military organizations differentiate among vertical system levels by using different terms, depending on the level under scrutiny. From highest to lowest to highest, this taxonomy is as shown in [Table 19-2](#).

The content of each of these terms decreases in “granularity” as one moves upward in the hierarchy. In other words, tactics are much more detailed, discrete, and narrowly focused than operations. Strategies are much more general and broad than operations, which themselves are more general than tactics.¹

Term	System Level
Grand strategies	Nations
Strategies	Unified commands (multi-service)
Operations	Larger units
Tactics	Small units

TABLE 19-2 System Levels

Non-military organizations don't normally make these distinctions, although they could—and perhaps should. Complex systems or organizations experience significant interdependencies among their internal components, the external environment, and other systems.

A Common Denominator

If one accepts that the concept of *strategy* embodies both vertical and horizontal dimensions, a real need for a common definition of the term emerges. Whether one calls it strategy, operations, or tactics, it answers the same underlying question: *How do we get from where we are to where we want to be?* Or, expressed another way, how do we achieve what we've set out to do?

Turning this question into a useful definition that suits both the variety of organizational types and the multiplicity of system levels, a “common denominator” definition of *strategy* might be:

How systems or individuals go about closing the gap between a current condition or position and a desired future state.

This definition is sufficiently inclusive to account for systems with multiple layers as well as different kinds of systems. It's not confined to military systems alone, nor is it exclusively centered on Marketing or Finance. Rather, it addresses both *means* (how) and *ends* (future state), regardless of the type or complexity of the system.

A Whole-System View

Means and ends don't exist in isolation. Every system having means and ends operates in some kind of environment. The nature of the environment—its economic, social, political, and technical characteristics—defines and delimits the resources and range of options a system can exercise in executing its strategy.

The relationship between a system and its environment naturally implies *decisions* about how to employ available resources in pursuit of the system's ends—in other words, in executing strategy. In the modern world, neither the

environment nor resource availability remains stable for long. The external environment is subject to a wide variety of variables, too. Consider, for example, the extreme fluctuations in international oil prices, the collapse of the U.S. sub-prime mortgage sector, and the failure of huge commercial banks. For most systems—commercial, government agency, or not-for-profit—such external factors, predictable and unpredictable alike, change their respective playing fields in dramatic and uncontrollable ways. Such turbulence continually generates situations requiring choices (decisions), any of which can affect outcomes or ends.

It's almost impossible—certainly impractical—to predict changes in the external environment with any confidence. The same might be true for the availability of resources. It is likewise impractical to preplan for an indeterminate number of contingencies that might happen. Such unpredictability drives a need for rapid, effective decisions, or reactions, during the execution of strategy—perhaps even the revision or replacement of the entire strategy. The point is that in the modern world, *strategy can never be static. It's inextricably linked to execution, and it must be continually reevaluated against the evolving conditions of an ever-changing environment.*

The OODA Loop

Perhaps the most influential development in the art of decision making in the past 30 years is the OODA loop (see [Fig. 19-2](#)). The name is an acronym for *observe, orient, decide, and act*. However, the OODA loop is considerably more robust than the mere sequential execution of the four steps the name implies.

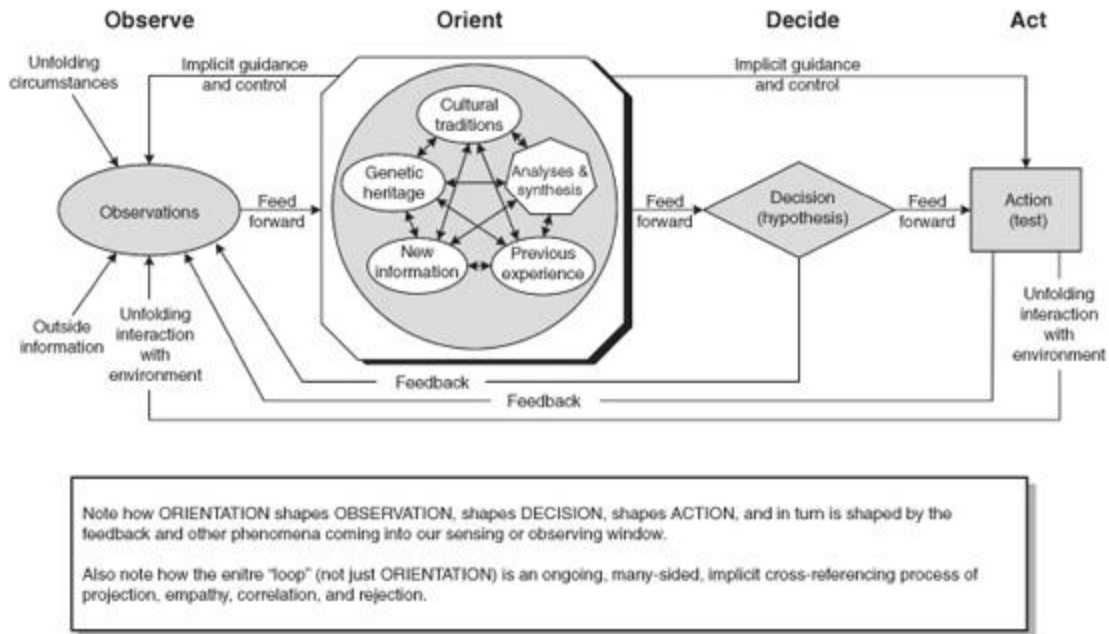


FIGURE 19-2 The OODA loop. (From Boyd, J. R. *The Essence of Winning and Losing*. 1996.)

In much the same way that the *Five Focusing Steps* (5FS) guide the management of system constraints in constraint theory (Goldratt, 1990), the OODA loop is a routine that facilitates rapid, effective decisions at all levels—tactical, operational, or strategic—of any kind of system, whether commercial, government agency, or not-for-profit.

The OODA loop is the conceptual brainchild of John R. Boyd, a U.S. Air Force colonel (1927–1997) who synthesized it from his personal experiences in air-to-air combat, energy-maneuverability theory, policy “battles” in the Pentagon, and extensive research into military history, strategy, and science. However, Boyd’s synthesis resulted in far more than the OODA loop alone, which is merely the most visible part of a larger system-level perspective on adjusting and evolving in an ever-changing world. (Coram, 2002; Hammond, 2001; Richards, 2004; Osinga, 2007; Safranski, 2008).

How does the OODA loop facilitate the development and deployment of strategy?

Strategy as a Journey

If one accepts the concept of strategy as summarized in [Fig. 19-3](#), a robust approach to decision making can mean the difference between success and failure in a rapidly changing environment. The first three stages of the *OODA loop*—*observe*, *orient*, and *decide*—are essential to the creation of strategy in the first place. The last stage—*act*—clearly applies to deployment of strategy. Nevertheless, it's called a "loop" for a reason—the first three stages also provide the means to detect and respond to the environmental changes that could rapidly render a strategy invalid.

Many companies use an annual strategic planning cycle, meaning that they have a predetermined yearly schedule for reviewing and updating their strategic plans. In other words, they set their strategy for at least a year, then don't formally revisit it until the same time next year. But how responsive is that practice to surprise, catastrophic events? How well would such a practice have served the commercial airlines after September 11, 2001, or commercial industries that depend on bank financing after September 2008? If strategy directs a journey from the current state to some desired future state, it's critical for it to be flexible enough to react immediately to such unexpected surprises. If you were navigating a ship across the ocean and discovered that you had been blown seriously off course, would you wait until the next strategic planning cycle to take corrective action? What if, for some reason, the destination had changed, even without a storm to blow you off course? Would you in any way delay resetting your direction? If not, why would anyone with responsibility for guiding organizations behave any differently?

- Strategy prescribes how to move from an existing condition to a desired future state
- Strategy applies (horizontally) to systems of all kinds:
 - Commercial
 - Government agency
 - Not-for-profit
- Strategy has whole-system implications (i.e., not confined to just a few functions)
 - Marketing
 - Finance
- Strategy has a vertical dimension as well as horizontal
- Strategy is the application of means (resources) to achieve ends (objectives)
- Strategies must consider the ever-changing nature of the environment in which systems function
- A constantly changing environment requires continual decisions to adjust or change strategy

FIGURE 19-3 Strategy as a journey.

Orientation and Observation

According to Boyd, the *orient* step is the most critical of all, despite the fact that it appears second in the sequence (Safranski, 2008) That's one reason why he made it more prominent (see [Fig. 19-2](#)) than any of the other steps. The orient step is the amalgamation or synthesis of the sum of our knowledge about ourselves, our system, values, customs, culture, experiences (heritage), and the environment (Osinga, 2007). One might oversimplify by saying that our orientation represents our worldview, hard won and tightly held. It's the lens through which we filter sensory inputs of things happening around us or, in other words, the *observations* we make in real time.²

The *orientation* step is the one in which a divergence from our expectations is detected. Part of our orientation is the paradigm (Kuhn, 1962) in which we live, the view of the world we create for ourselves based on the factors previously mentioned. These factors all conspire to form our

assumptions about the way we think things happen (or should happen). When we observe phenomena or events that don't fit into our orientation, we have what Boyd referred to as a *mismatch*. The existence of this mismatch is determined when we analyze and synthesize our observations with the basis of our orientation or paradigm. In other words, we examine what *is* happening in light of what we expect *should* be happening. This continual analysis-synthesis process is an integral part of maintaining a robust current orientation.

How does *observation* happen? Sometimes, as in the case of 9/11 or the sub-prime mortgage meltdown, events are thrust upon us in ways that we can't ignore. However, sharp system leaders actively look for changes in the environment and evaluate what effect their observations might have on their orientation—in other words, what mismatches might be emerging. The more this active observation is practiced—and the observations synthesized—the more sensitive one eventually becomes to small changes, which may be indicators of more dramatic changes yet to come. This has relevance to competitive advantage, which will be discussed in more detail shortly.

As [Fig. 19-2](#) indicates, observations include new outside information, such as research or technology breakthroughs. Unfolding circumstances include the entry of new competitors into the market, new laws or regulations, or world events such as skyrocketing crude oil prices, increased activity of Somali pirates in the Indian Ocean, financial chaos in one sector of the economy, or other international geopolitical developments. Unfolding action with the environment specifically refers to the environmental effects of actions the system might take—the other side of the equation from the impact of environmental changes on the system. Implicit guidance and control (at the top-left in [Fig. 19-2](#)) represents the changes in a system leader's observations based on the synthesis of new

information, even before decisions or actions are contemplated.

Decision and Action

Completion of the orientation step implies that a mismatch or gap between reality and expectations has been identified. The next step would seem to be to *decide* what to do about it. The decision step in the OODA loop may be deliberate or intuitive. In complex situations, when the decision maker isn't intimately familiar with the environment or the possible options, this step is likely to require deliberation: "We know that things are not the way they should be—now what should we *do* about it?" A more formal or structured decision process might ensue.

However, if one's knowledge of the system and its environment is comprehensive (usually born of deep experience), it may be intuitively obvious what needs to be done. In this case, decision makers often proceed directly to action. This is reflected in the upper-right part of [Fig. 19-2](#) (implicit guidance and control).

Even if decision making is more deliberate, available options are often logically tested—that is, compared to reality and their potential outcomes assessed—before proceeding to the action stage. This "hypothesis testing" is reflected in [Fig. 19-2](#) in the feedback loop between "Decisions" and "Observations." The purpose of this testing is to help reduce the impact of uncertainty on a decision among several options.

Inevitably, however, even with the hypothesis-testing feedback loop, the ultimate end of the OODA process is an action of some kind. And because action inevitably influences the environment in some way—after all, that was its purpose in the first place—the process begins all over again with observing to assess the action's impact. This in

turn begets a second iteration of the orientation step to determine how much impact the action had, whether it changed reality in the desired direction, and by how much. The size of the mismatch that results from this second orientation leads to another decision and subsequent action. And the process continues until the ultimate goal of the system is attained.

“Pro-Acting” Rather than Reacting

Superficially, it might seem that the OODA loop is reactive. However, Boyd’s contention was that controlling an emerging situation was far preferable than reacting. Consequently, his prescription for using the OODA loop was anything but passive. He was highly motivated to “stir the pot”—to use the OODA process to *create* mismatches, especially in the perception of adversaries. In this respect, he recommended being pro-active, rather than reactive.

However, rational decision making and action depends on a conscious awareness of these four steps: *observe*, *orient*, *decide*, and *act*. In reality, most people actually do something like this, but they do it unconsciously or intuitively. They’re usually unaware that they’re doing it, which means that they are less likely to “keep the pressure on.” Without consciousness about the OODA process, like the fabled hare they’re likely to take a nap alongside the road while the tortoise passes them by.

Fast OODA Loop Cycles

Boyd went even further with the pro-active OODA concept. He contended that if one could cycle through these four steps faster than one’s adversary could, a competitive advantage would begin to open up. The non-OODA

practitioner would always be at least one cycle behind the OODA user. Moreover, if the OODA user could somehow complete two or more cycles in the time the adversary took to finish one, it would sow confusion in the opponent's camp. In battle (the context for which Boyd created the OODA loop), this ultimately results in panic, knee-jerk (wrong) reactions, and eventual collapse of the opponent.

The effect is not materially different in business settings. Witness, for example, the introduction of high-technology innovations by the Japanese for nearly two decades. It was commonly recognized that while the world's markets were enamored of their latest, greatest product introduction (first the Walkman, then CDs, then digital cameras, then compact video devices, then DVDs and MP3 players, etc.), the Japanese were hard at work on the "next big thing." The rest of the world was always at least one step behind.

Boyd himself provided the original, quintessential example of the fast-cycle OODA loop. As a U.S. Air Force fighter weapons instructor in the 1950s, he made a standing offer to all pilots: He would beat his opponent in 40 seconds or pay them \$40. In eight years, no one was ever able to collect the \$40 (Coram, 2002). The reason was that he was always able to execute what amounted to a near-instantaneous OODA cycle faster than any of his opponents could.³

Summarizing Boyd

Let's quickly review what we've just covered.

- The OODA loop describes a process of observing, synthesizing those observations (orientation), deciding what to do because of the synthesis, and acting on that decision.

- Although all systems go through this OODA process, most are completely oblivious to the fact that they're doing it.
- The OODA loop was originally conceived as a way of mentally managing combat engagements to achieve victory, but its applicability in the development and deployment of strategy has yet to be fully realized.
- The OODA loop appears, on the surface, to be reactive to changes in the environment; however, a deft practitioner can use it proactively to shape the environment or competitive arena to his or her own advantage.
- The ability to cycle through the OODA loop multiple times while others do so only once can provide an insurmountable competitive advantage.

Armed with this knowledge of systems and the OODA loop, leaders can enjoy a substantive potential advantage over others (and the environment) in achieving their systems' goals. However, this advantage remains exclusively potential without discrete tools with which to execute the OODA loop.

The Logical Thinking Process

Concepts such as the OODA loop are eminently useful but sometimes difficult to translate to practical application without some kind of tool to bridge the gap between the conceptual and the practical. Fortunately, the appropriate tool for applying the OODA loop strategically is readily available: The *Logical Thinking Process* (LTP).⁴

The LTP is an outgrowth of the evolution of TOC. Originally conceived as a production scheduling and management methodology called "Drum-Buffer-Rope" (Goldratt, 1990), in

the late 1980s and early 1990s TOC outgrew its former production-oriented boundaries and spread into the broader category of systems. One of the first such forays was the thinking process. When it became obvious that resolving production bottlenecks alone didn't always produce a more successful company, Goldratt needed another solution. He conceived the thinking process to address the application of his 5FS (Goldratt, 1990) when system-level constraints were not production bottlenecks—when the factor limiting overall system success lay in non-production areas.

This was a critical breakthrough because it raised the whole idea of constraint theory to a system concept, rather than just being a production methodology alone. The thinking process afforded a means to examine systems of any kind, not just production companies, and identify the one factor limiting the system the most in its mission to achieve its goal.

Originally composed of five logic trees or tools,⁵ the thinking process represented a simple application of the scientific method to the challenge of complex system problem solving: what's the problem (what to change), what do we do about it (what to change to), and how do we do it (make the change happen)? For the first time, the thinking process offered a concise, direct way to logically analyze whole systems composed of myriad complex interactions and do so rapidly. Moreover, it also allowed for "hypothesis testing" without extensive real-world experimentation to verify the validity of proposed changes. In addition, what it also did that no other problem-solving methodology did was to include a solution implementation "module"—the prerequisite and transition trees. In other words, a complete package. [Figure 19-4](#) illustrates the conceptual flow of the thinking process as originally conceived by Goldratt.

Over the intervening years since Goldratt introduced the thinking process, the trees and their application have

evolved and been refined. Although the process was originally intended to solve complex problems by identifying system constraints and facilitating ways to break them, it was inevitable that other applications would emerge. One of these was the use of the thinking process for strategy development and deployment (Dettmer, 2003). However, applying the thinking process for strategy development purposes requires some modification of both the trees and their sequence. To distinguish these evolutions from the original thinking process, the term “logical thinking process” is used hereafter.

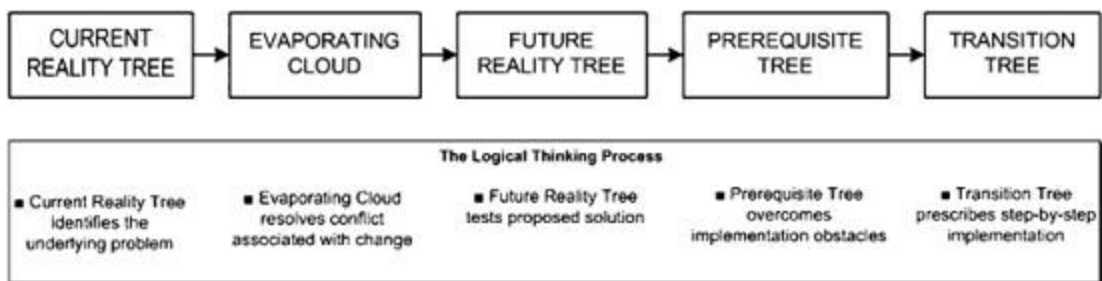


FIGURE 19-4 The Logical Thinking Process.

The Intermediate Objectives Map

The most significant modification to the LTP for strategy development is the insertion of a new type of tree—the *Intermediate Objectives (IO) Map*—at the beginning of the process (Dettmer, 2007). The IO Map is critical to the strategic application. In fact, without it, the remainder of the LTP is nearly useless for strategy development.⁶

The IO Map is a relatively simple structure, but actually putting one together requires some dedicated thinking. [Figure 19-5](#) shows a conceptual version of the IO Map. An actual IO Map may be found in [Fig. 19-11](#) at the end of this chapter.

The goal indicated at the top of the IO Map is the ultimate outcome for which the system strives. In a for-profit commercial company, this is usually maximum profit. In not-for-profit organizations, such as charities or hospitals, the goal is usually some favorable contribution to society. Goals of government agencies are likewise not profit-oriented, but rather seek the successful provision of some beneficial service to the general population.

Every goal is typically achieved by realizing a set of *critical success factors* (CSFs). These CSFs are terminal outcomes, or results. They're considered critical because they're indispensable to attainment of the goal. In any system, and for any goal, very few CSFs are normally required to declare goal attainment. For most systems, they would number no more than three to five. CSFs represent very high-level outcomes. They are usually somewhat generic to the category of the system under discussion. For example, the CSF for any profit-oriented company would be quite similar, differing primarily only in degree of emphasis. If the goal of a commercial company is to maximize profits, there are really only three CSFs: increased Throughput, minimized Inventory, and controlled Operating Expenses (see [Fig. 19-6](#)).

Notice that none of these differs, whether the company is an automobile manufacturer or an insurance company. If these CSFs are realized, then the inevitable outcome is a company that has maximized profitability.⁷ Where do the specific details of company activities (processes, products, competitive factors, etc.) fall? They lie beneath the level of the CSFs themselves, in what [Fig. 19-5](#) depicts as *necessary conditions*. It is at the necessary condition level that the unique picture of a particular organization emerges. [Figure 19-7](#) shows how this might look for a typical manufacturing company.

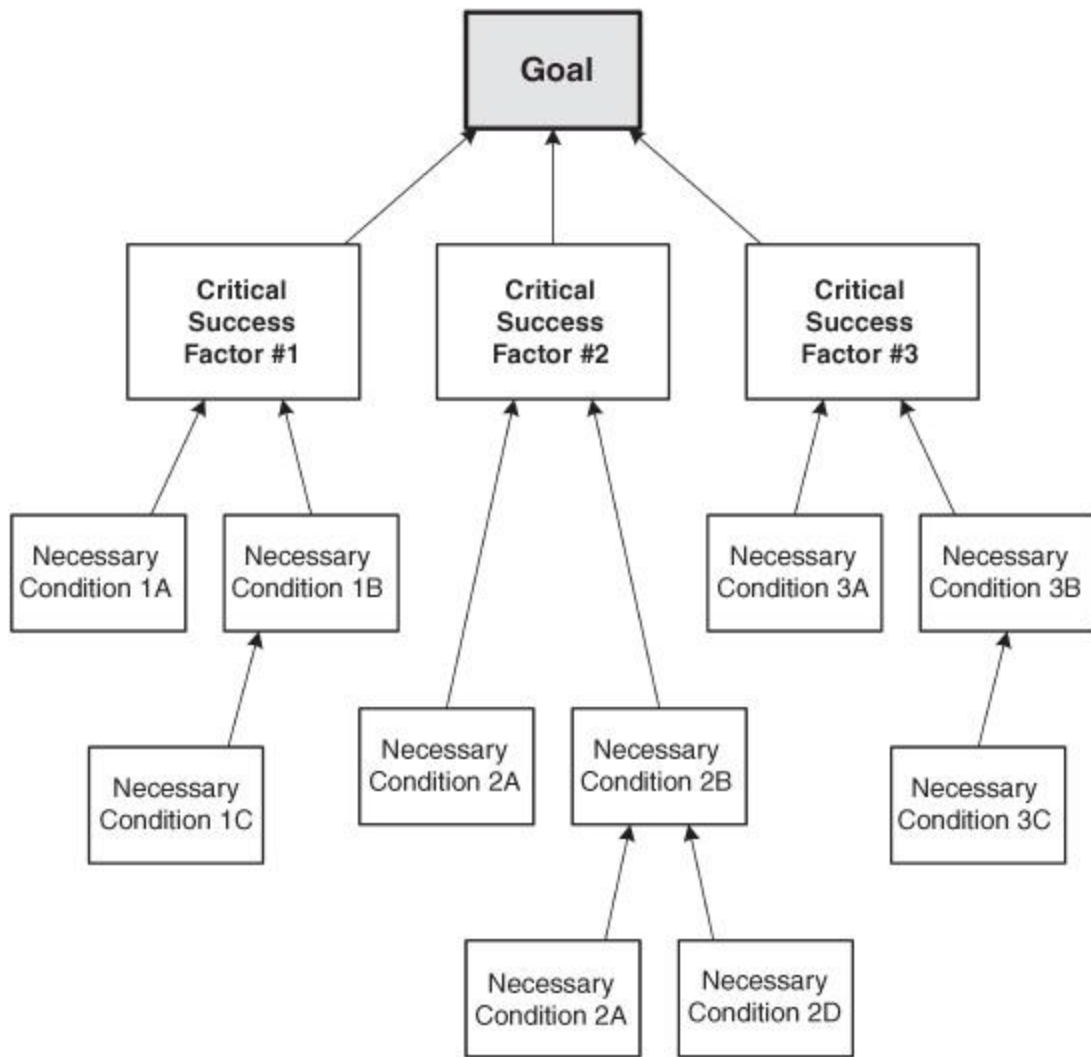


FIGURE 19-5 Intermediate objectives map.

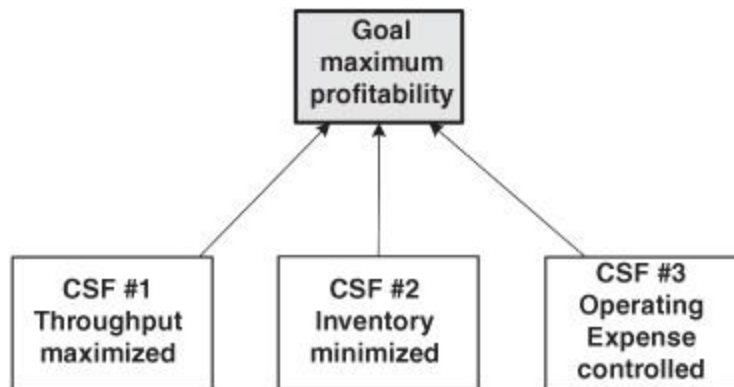


FIGURE 19-6 Goal and critical success factors (commercial company).

The CSFs of a not-for-profit organization or government agency would be somewhat different from those of a commercial company. For one thing, neither usually measures its Throughput financially, but rather in terms of whatever non-pecuniary benefit the organization is in business to provide for society. Minimum Inventory and controlled Operating Expense might certainly be relevant, however.

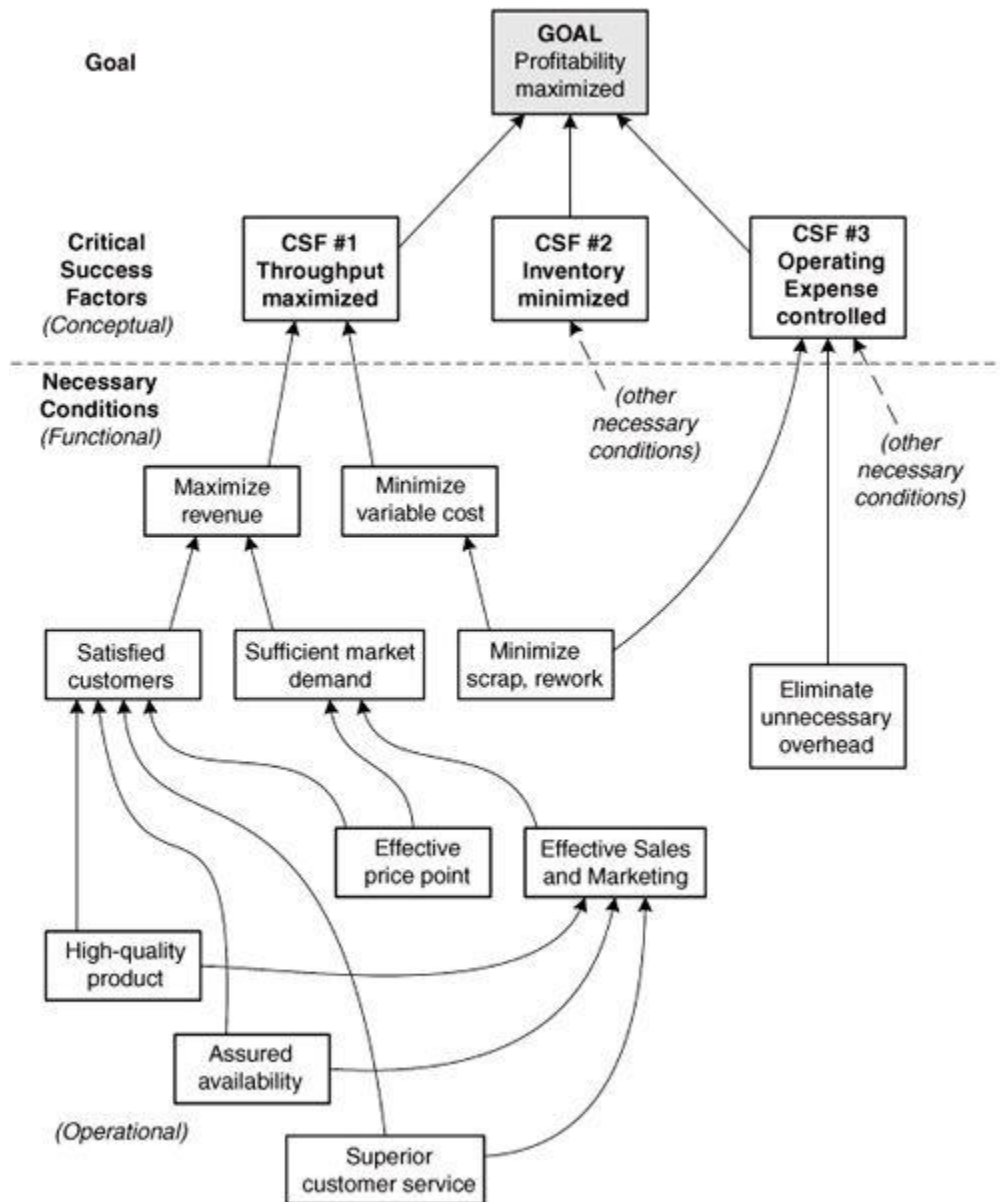


FIGURE 19-7 IO Map (partial)—commercial company.

The question of where to put such non-negotiable requirements such as adherence to the law, compliance with regulations, or environmental responsibility inevitably comes up. None of these factors, and others comparable to them, directly affect profitability, so they clearly don't fit as critical success factors. However, they usually do serve to define the behaviors associated with fulfilling them. In other words, their proper place is as necessary conditions for the

generation of Throughput, the reduction of Inventory, or the control of Operating Expense. This positions them at least three layers down in any IO Map, and probably even lower.

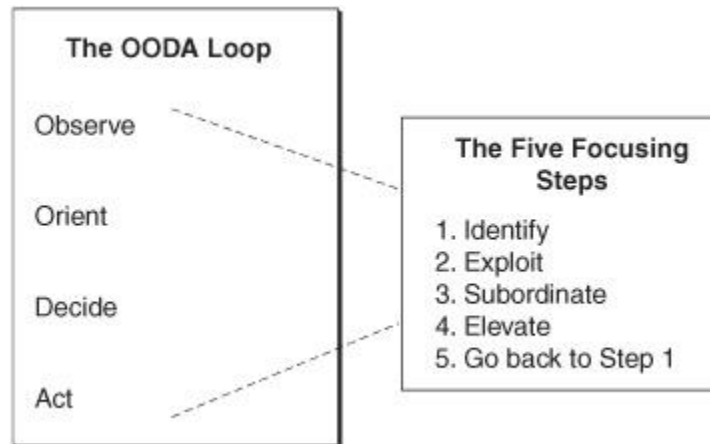


FIGURE 19-8 OODA loop and the Five Focusing Steps.

How far down should the IO Map be “drilled?” For constructing a subsequent *Current Reality Tree* (CRT), it’s not necessary to go much below the CSF and perhaps one or two layers of necessary conditions. However, for resolution of conflicts that might develop in using the LTP for either strategy development or for complex problem solving, it might be advisable to penetrate down five or six layers.

When the IO Map is completed, it provides two crucial ingredients for the successful application of the rest of the LTP. First, it clearly delineates the discrete activities and outcomes required to ensure achievement of the system goal (without regard to what is actually happening at the moment). Second, it provides the basis for consensus among everyone within the system—executives, managers, and specialized employees alike—on what they should be doing to support one another in a coordinated way. This might be called a “unified vision” of where the company is going and what’s required to get there.

Constraint Management Model: A Synthesis of TOC and the OODA Loop

The 5FS, the heart and soul of constraint theory, constitute the guiding framework for real system improvement. The OODA loop represents an articulated model for a true cybernetic system—one that is not only capable of self-improvement, but self-determination of direction as well.⁸ There is an implicit relationship between the two (see [Fig. 19-8](#)).

The 5FS are inherently a subset of the OODA loop. Identification of system constraints requires observation and orientation (the first two steps in the OODA loop). Exploitation, subordination, and elevation are all elements of the decision step in the OODA loop. The actions to follow the prescriptions of the 5FS are the same as the final step of the OODA loop. Both employ a feedback process to begin the cycle again. What makes the OODA loop more generic than the 5FS is its applicability to situations in system operations that don't involve identifying and breaking constraints or dedicated system improvement effort.

Boyd originally conceived the OODA loop to help manage tactical operations. The O-O-D-A (and repeat) cycle is inherent in activities as narrowly focused as driving a car safely on a winding road, or as broad as steering the progress of a corporation into its future. However, it's this last, broader perspective with which we're concerned when we talk about strategy.

If we accept the idea that developing and deploying strategy is an expression of the OODA loop, the question that naturally follows is, "How do we go about doing this?" This is where the LTP offers an ideal solution. The combination of the OODA loop and the LTP produces the *Constraint Management Model* (CMM) for strategy

development and deployment (Dettmer, 2003). It's so named because the LTP was derived from the effort to apply TOC to whole systems, and in using the LTP to develop and deploy strategy the management of constraints is a natural byproduct. In other words, you can't effectively execute whatever strategy you might develop without identifying and breaking your existing system constraints. [Figure 19-9](#) illustrates the CMM.

The CMM is, itself, a seven-step cyclical process.

Step 1. Define the paradigm. The first step in any strategy development process should be to define the system, its goal and CSFs, and the characteristics of the environment in which it operates. This is where the first three levels of the IO Map are developed. Besides some serious conceptual thinking, this naturally requires both internal and external observations to be made—the first step in the OODA loop.

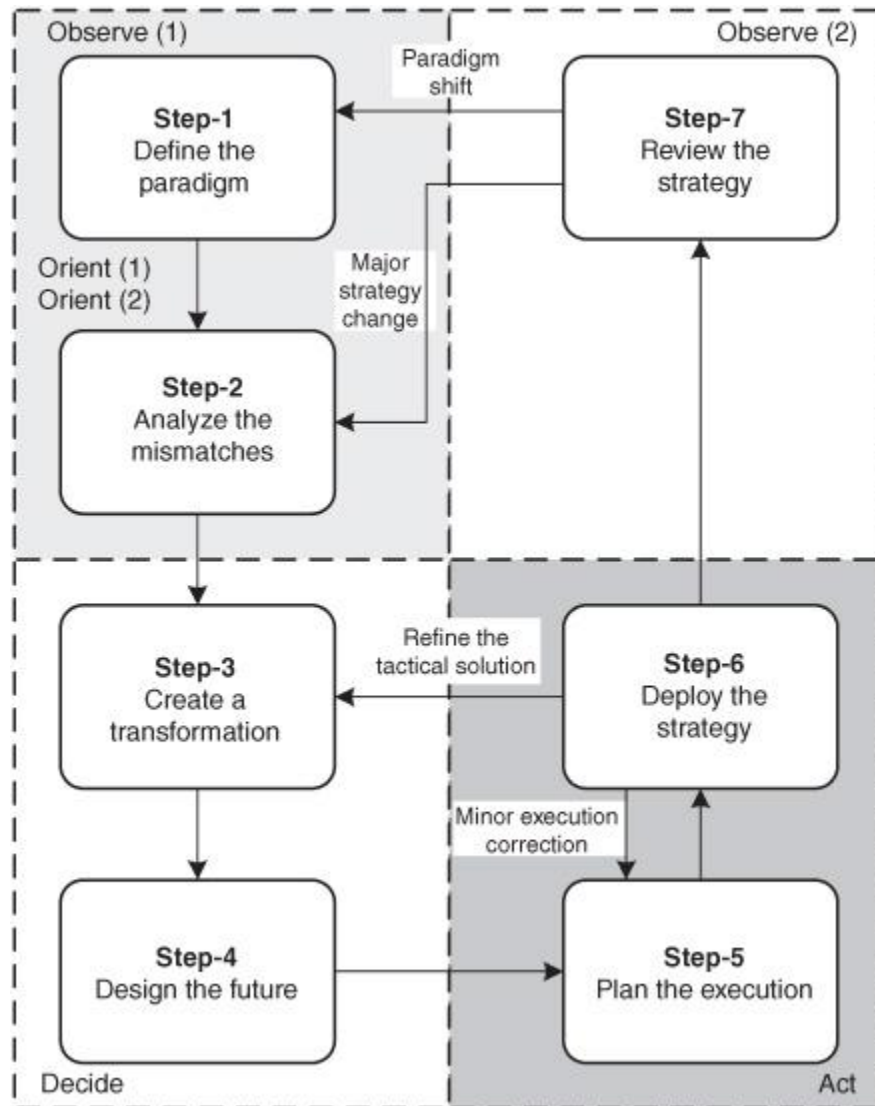


FIGURE 19-9 The constraint management model. (From Dettmer, H. W. 2003. *Strategic Navigation: A Systems Approach to Business Strategy*. Milwaukee, WI: ASQ Quality Press.)

Step 2. Analyze the mismatches. Once the system and its operating environment are defined and observations of the current situation made, it's time to synthesize what *should be* happening with what *actually is* happening. This synthesis is the essence of Boyd's *orientation* step in the OODA loop. The product of this synthesis is one or more gaps, or what Boyd referred to as "mismatches." In this

case, the mismatch is between reality and our expectations. The size and scope of such gaps are specifically articulated. Inevitably, a system's current constraint will be found somewhere within the identified mismatches.

Step 3. Create a transformation. This is essentially a "brainstorming" step. It's the point in the process where creativity is required—thinking "outside the box" to create breakthrough ideas. Such ideas must be created before any decisions about what to do can be made. "Creation" is an inspirational or inventive activity. There are several widely used idea-generation methods, such as TRIZ (Rantanen and Domb, 2002), that can contribute breakthroughs in thinking needed to close the gaps discovered in Step 2.

Step 4. Design the future. Once a breakthrough idea (or more than one) is created to close the gap defined in Step 2, it must be integrated into a whole-system plan that includes not just the changes to close the gap, but the continuing operations that had no mismatches associated with them. Hypothesis testing, whether in the form of a simulation, prototype, or just a logical verification, verifies the efficacy of various alternatives, from which one or more are selected. This is the essence of the decision step in the OODA loop.

Step 5. Plan the execution. Once the decision is made, an execution plan should be formulated, since "the devil is in the details." Resources, accountabilities, timelines, and measures of success are established in execution planning. (If this is beginning to sound like a project, it's because it *is!*) An execution plan represents the "front end" of the OODA loop's *act* step.

Step 6. Deploy the strategy. This is the conclusion of the *act* step. How long the execution actually takes will depend on the nature of the activities planned. Strategies are typically longer-range than business plans or tactical actions. Time horizons are often measured in years. However, the completion of Step 5 makes managing

deployment better structured and easier to monitor. Moreover, as the inevitable surprises, deviations, or unexpected variations occur in execution, the plan can be expeditiously corrected to accommodate them. This is the second half of the OODA loop *act* step.

Step 7. Review the strategy. Presuming that no major breakdowns in strategy deployment occur, the only remaining task is to evaluate the strategy's overall effectiveness. This obviously brings us back to the OODA loop's first step again—*observe*. This time, however, we're not looking for deviations in deployment. We're determining whether the overall strategy we developed in Step 4 is really producing the results we want and expect.

Step 7 includes two feedback links. The more common one connects to Step 2 again (analyze the mismatches). Working with our previously defined paradigm and expectations (established the first time through the OODA loop in Step 1), we compare the second round of observations with our original expectations.⁹ Have the gaps identified earlier narrowed or even closed altogether? If not, or if they're not closing quickly enough to suit us, we must reevaluate our strategy and adjust it as necessary. Even if the gaps have closed, a proactive application of the OODA loop requires that we identify and develop "the next big thing" in our chosen field of operation. For example, Sony didn't sit on their Diskman[®] audio players or Trinitron[®] televisions after they stormed the market with them. They immediately began working on an MP3 player and a flat-screen video display. *That's* being proactive. The second, and less obvious, feedback loop takes us through Step 1 again. This is likely to happen much less frequently than the other feedback loop. This particular loop implies that a complete re-examination (and perhaps redetermination) of goals, critical success factors, and the external environment is required. In other words, it's possible that dramatic

change in the external environment of such magnitude has precipitated a complete redesign of strategy. What kind of event might this be? How about an economic depression or some catastrophic event such as a world war? Take Toyota, for example (Holley, 1997). Originally (before World War II), it was a manufacturer of textile machines. By the end of that war, its surviving manufacturing base had been completely converted to *automotive* vehicles, at the insistence of the Japanese Imperial Army. That was a conversion forced on Toyota by circumstances. However, by 1997 Toyota was anticipating that within 100 years the automobile segment of their business would constitute no more than 10 percent of the total. The rest would be in low-cost prefabricated housing and information systems. *These* are strategic shifts—proactive ones.

The Role of the LTP in the CMM

How does the LTP fit in with the CMM? The preceding description of the CMM fairly begs for a structured tool to make Steps 1 through 5 happen. That tool is the LTP. [Figure 19-10](#) shows how the LTP energizes the CMM.

The IO Map is used to establish the benchmark of expected or desired performance. For an organization that already understands that it's not yet where it wants to be, the articulation of the goal and CSFs in the IO establish a "stake in the ground"—the destination marker that determines where the organization wants to be at the end of the strategy's time horizon. Supporting necessary conditions represent the high-level functional milestones that must be achieved to reach the goal. Inherent in the development of the IO Map are research, observations, and information gathered about the external environment.

With the IO Map as the entering argument (desirable state), a CRT¹⁰ is constructed to depict the relationship

between reality and the end results depicted in the IO Map. The resulting gaps are reflected as *undesirable effects* (UDEs). The construction of the body of the tree, down to the critical root causes, embodies the synthesis (or orientation) of newly acquired knowledge about the external environment with experience, expertise, custom, tradition, etc.—the existing paradigm, if you will. The CRT produces the logical causes of the gaps (UDEs), without regard to whether they are politically acceptable to consider changing.

Especially in the latter situation, the transformation created in Step 3 is facilitated by the use of *Evaporating Clouds* (ECs), which are specifically designed to resolve intractable dilemmas such as political feasibility. The output of the ECs, and the beginning of this transformation process, is one or more injections that represent breakthrough ideas. These ideas become *initiatives*, or new projects that will provide the impetus to move the organization from where it is to where it wants to be. Some of these initiatives (changes) will undoubtedly be externally focused. Others will be inwardly directed.

The *Future Reality Tree* (FRT) takes these initiatives, or ideas, and logically structures them to verify that, in fact, they will move the organization toward its ultimate goal. The reflection of that movement is in the narrowing, or complete closure, of the gaps identified in Step 2. This narrowing/closure is represented as a *desired effect* (DE) in the FRT. Besides logically verifying that the initiatives created will, in fact, advance the organization toward its ultimate goal, the FRT will include the “ferreting out” of negative branches—those conditions under which the whole strategy deployment (or key aspects of it) might be derailed. The “trimming” of these negative branches becomes contingency plans. The completed FRT, with trimmed negative branches, is the organization’s strategy. The FRT injections are the strategic initiatives, programs,

projects, etc. required to impel the organization toward its goal.

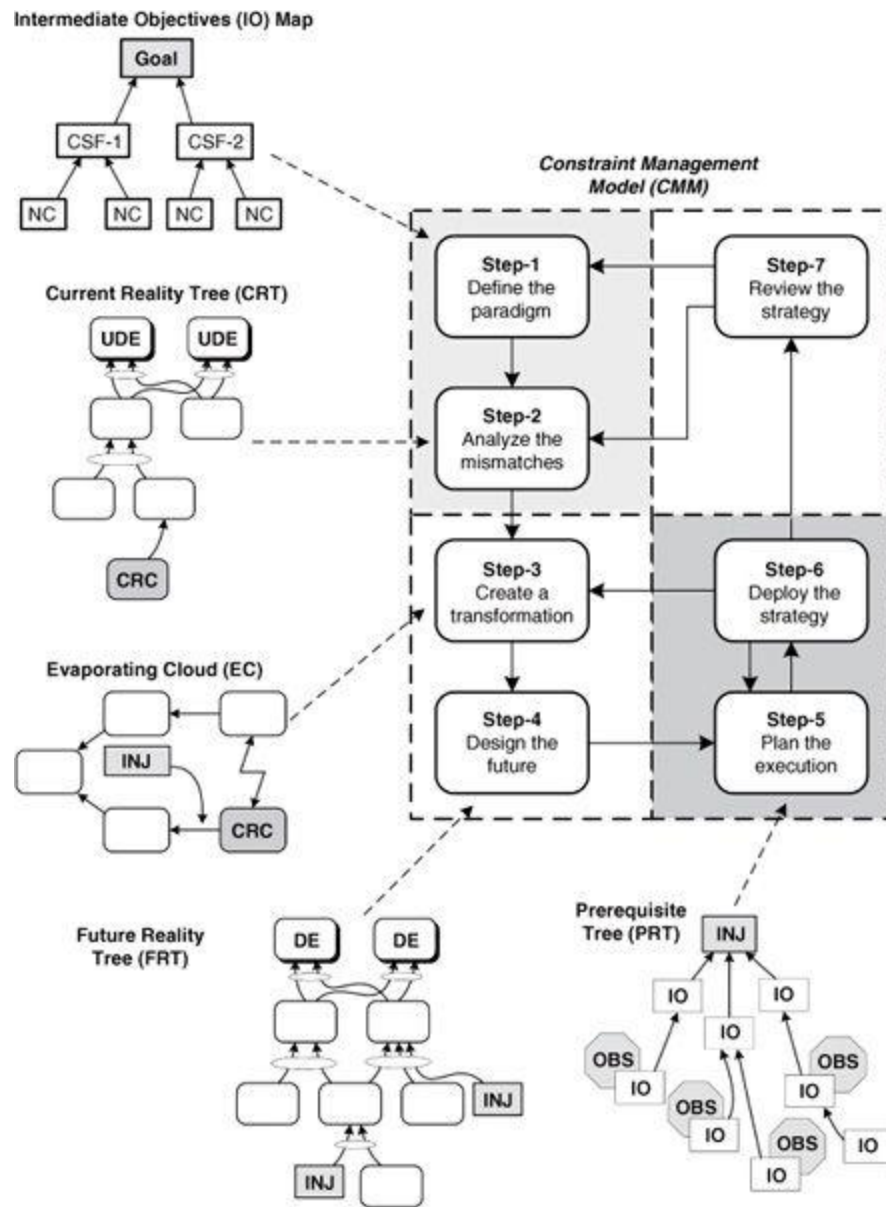


FIGURE 19-10 The logical thinking process and the constraint management model.

Once the strategy is developed in Step 4 as the second part of the *decide* stage in the OODA loop, the *action* stage naturally follows. Step 5 is the detailed execution planning. Each of the injections, or initiatives, defined and verified in

the FRT (Step 5) is “fleshed out” in a *Prerequisite Tree* (PRT). Obstacles are overcome and important milestones and sequential/parallel tasks are identified. The resulting PRT forms the basis of a project plan—a project activity network—that can be managed using *Critical Chain Project Management* (CCPM). The consolidation of all PRTs into multi-project CCPM becomes the organization executive’s tool for managing the overall long-term deployment of the strategy.

What about Steps 6 and 7?

The natural question at this point is, “But what about Steps 6 and 7 of the CMM?” The answer is that at the conclusion of Step 5, the role of the LTP ends. Strategy deployment (Step 6) is an ongoing leadership responsibility. Effective executives use a variety of tools and techniques to shepherd a deployment along. If the execution planning in Step 5 included conversion of PRTs to a CCPM schedule, then one of the obvious TOC-related tools a leader might use at this point is *Buffer Management* (BM).

Step 7 is an executive function, too. It requires a conscious, deliberate effort to repeat the *observe* step of the OODA loop again with the objective of identifying failure of the strategy to deliver the intended results and the reason for that failure. In many, perhaps most, cases such failure has less to do with the inadequacy of the strategy than it does a rapid, possibly catastrophic shift in the environment. How many perfectly good strategies do you think might have been rendered ineffective by the 9-11 terrorist attacks in 2001, or the collapse of the U.S. economy in 2008? Even if the triggers are not quite so dramatic, such environmental changes can prompt a need to reevaluate and adjust strategies—or even replace them

altogether. And so begins the second iteration of the OODA loop with a return to the IO Map and CRT.

Summary and Conclusion

Formal strategic planning in business dates back only to about 1965, although the development and employment of strategy have been practiced since the days of Sun Tzu some 2500 years ago. In contemplating strategy, there are some worthwhile points to keep in mind.

- Distinguish between the development of strategy and a strategic plan. The latter is no more than the capture in some written form of the former. Strategy development, not the written plan, should be the primary focus.
- For businesses, strategy is about far more than just marketing and sales. It's concerned with the long-term attainment of the organization's goal. If that organization is a commercial company, Marketing and Sales will be but one part of that effort.
- Organizations live or die as complete integrated systems, existing in an external environment that imposes conditions, including competition, on the activities of the system. Effective strategy must consider both the internal activities and the external environmental factors.

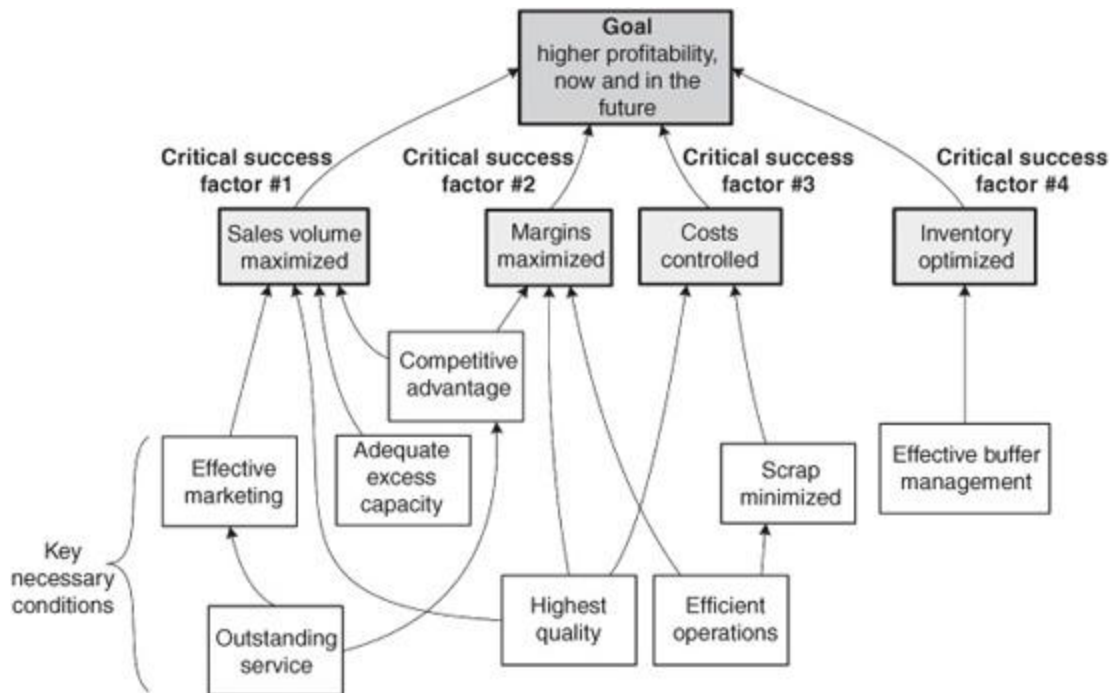


FIGURE 19-11 AllForm Welding Company strategic intermediate objectives map.

- The OODA loop developed by Boyd provides an excellent foundation for managing the development and evolution of strategy over the foreseeable time horizon of an organization. (It should be emphasized, however, that the OODA loop is only one small but important part of Boyd's contributions to systemic thinking. The sources on Boyd listed in the references are all highly recommended reading.)
- The LTP is perhaps the most powerful system-level policy analysis tool ever conceived. Strategy development and refinement is very much concerned with policy analysis, since strategic prescriptions inevitably take the form of policies to some degree. Consequently, the use of the LTP as a strategy development and deployment tool can't be reinforced too strongly.

- Merging the framework provided by the OODA loop with the trees of the LTP provide a “power boost” for organizations of any stripe—commercial, not-for-profit, or government agency—in helping them achieve their goals. If such organizations exist in a “zero sum” environment (a gain for them is a loss for some other group), this kind of assist can spell the difference between success and failure.

References

Athey, T. H. 1982. *Systematic Systems Approach: An Integrated Method for Solving System Problems*. Upper Saddle River, NJ: Prentice-Hall.

Coram, R 2002. *Boyd: The Fighter Pilot Who Changed the Art of War*. New York: Little, Brown & Co.

Dettmer, H. W. 2003. *Strategic Navigation: A Systems Approach to Business Strategy*. Milwaukee, WI: ASQ Quality Press.

Dettmer, H. W. 2007. *The Logical Thinking Process: A Systems Approach to Complex Problem Solving*. Milwaukee, WI: ASQ Quality Press,

Goldratt, E.M. 1990. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Great Barrington, MA: The North River Press.

Hammond, G. T. 2001. *The Mind of War: John Boyd and American Security*. Washington, DC: The Smithsonian Institution Press.

Holley, D. 1997. "Toyota heads down a new road," *Los Angeles Times*, March 16.

Kuhn, T. 1962. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press.

Osinga, F. P. B. 2007. *Science, Strategy and War: The Strategic Theory of John Boyd*. New York: Routledge.

Porter, M. E. 1985. *Competitive Advantage*. New York: The Free Press.

Ramel, G. Gordon's Flea Page. Siphonaptera: A nursery rhyme, dating back to the 1800s. <http://www.earthlife.net/insects/siphonap.html>.

Rantanen, K. and Domb, E. 2002. *Simplified TRIZ: New Problem-Solving Applications for Engineers & Manufacturing Professionals*. Boca Raton, FL: St. Lucie Press.

Richards, C. 2004. *Certain to Win: The Strategy of John Boyd Applied to Business*. Philadelphia, PA: Xlibris Corporation.

Safranski, M., Ed. 2008. *The John Boyd Roundtable: Debating Science, Strategy and War*. Ann Arbor, MI: Nimble Books LLC.

About the Author

William Dettmer is senior partner at Goal Systems International, providing consulting and training on established applications of constraint management tools in both manufacturing and services with Fortune 500 and other companies. He has developed new applications for

constraint theory, principles, and tools. Dettmer has deep experience in logistics, project planning and execution, and contracting/procurement and has had direct responsibility for project management, logistics planning, government contracting, system design, financial management, productivity improvement, idea generation, team building, strategic planning, and customer-supplier relations. He is the author of seven books on constraint management and system improvement.

CHAPTER 20

The Layers of Resistance—The Buy-In Process According to TOC

Efrat Goldratt-Ashlag

Introduction

Sitting in a crowded airport lounge not long ago, I overheard a discussion between two men regarding a proposed change in their organization. The first man was making a real effort to convince his colleague to go along with the change. The colleague was clearly not thrilled about the idea, and began raising objection after objection. As soon as the first man addressed those concerns, his colleague was either ready with a new objection or, worse, insisted on rehashing a problem the two had already discussed. As the men grew more and more irritated with one another, all I could think of was how I wished these guys were familiar with the Layers of Resistance—*that* might have given them a chance to get somewhere instead of going around in circles.

When we recognize that a change should be made, we often realize that we cannot pull it off without someone else's permission and/or collaboration. Thinking about bringing another party on board tends to make us somewhat apprehensive. Not only because of the time and effort it is going to take, but mostly because we can't be sure that these efforts will pay off; getting buy-in is not a trivial task. So, we prepare our arguments (or don't), take a deep breath, and tell the other party all about our fantastic

idea. Sometimes it works and they get excited, and sometimes it doesn't and they leave less than enthused. Resistance comes in many forms: We might encounter a flat-out NO, or get caught in the cycle of objection and reassurance like those folks at the airport. Even a repetitive "let me think about it" can be a type of resistance, and there are many more. The result is still the same: We do not yet have the approval or collaboration we need in order to move on.

Our natural reaction on such occasions is to get all worked up and blame the other party for being indifferent or stubborn or even stupid (Goldratt, 2009). After all, they are the ones who failed (miserably, we might like to tell ourselves) to see the need for our change. The literature on the subject also focuses on the other parties' reasons for resisting change, stating causes such as personality traits (e.g., intolerance for ambiguity, need for control), inertia, promoting or protecting one's self-interest, and more. If we pause for a minute to think about what these causes mean, we can see that the literature has a lot in common with our natural reaction—both imply that the person who resists the change is the "bad guy" in the situation. TOC takes a very different stand in that matter. What TOC suggests is that instead of blaming the other party, the person proposing the change should be accountable for thoroughly planning and presenting the change.

Copyright © 2010 by Efrat Goldratt-Ashlag.

First of all, let us assume that we are talking about a win-win change, one that benefits all parties involved. Too often we come up with the most creative justifications for demanding that others give up their needs so that we can get ours met. We are pushing for a win-lose change. If we expect to "win" at the expense of the other side, we are practically asking for resistance—and shouldn't be surprised when we get it. Win-lose solutions are hard to sell, and even

if we have the power to enforce them, we cannot expect our partners to collaborate happily. In this chapter, we focus only on win-win changes.

At first glance, it seems that win-win changes should be easy to sell. After all, if everyone wins, why would anyone object? Win-win changes should practically sell themselves. In reality, however, this turns out to be false. People do object to win-win solutions, and often for very good reasons. For instance, they may not be clear on *how* they win, exactly (or, shall we say, we haven't outlined their benefits clearly enough), they may have concerns that we might have overlooked, they might believe this change will not "stick" and want to preserve their energy for more worthwhile efforts (as excited as we are about this change, have we really thought of how to integrate it fully?), and so on. Today's world presents people with abundant opportunities to make changes in all areas of life. In order to make sure that they look after their best interest and use their resources for efforts that will pay off, it stands to reason that people will approach change with various degrees of caution. If we would like to implement a change that requires their collaboration, then it is up to us to buy them in.

Some of us are excellent salespeople when it comes to getting people on board, and some of us are less "talented" in that area. We have all initiated and implemented successful changes in the past, but we have probably failed too—we have tried to get others to collaborate and we have gotten stuck. The question is, when we get stuck, is there something we can do about it? Is there a way for us to uncover the other party's concerns and address them properly? Or, if we anticipate difficulty in getting buy-in for a certain change, can we tackle it in advance? Can we systematically line up our arguments, so that we have a better chance of getting people to collaborate with us? The

TOC “Layers of Resistance” may offer significant insights into these questions.

The Layers of Resistance to Change

The Layers of Resistance to change originate from the TOC basic questions of change (Goldratt, 1984).

1. What to Change? (What is the problem we are attempting to address?)
2. What to Change to? (What is our solution to this problem?)
3. How to Cause the Change? (How to implement the solution?)

Taken together, these three questions represent the buy-in effort in a nutshell. Yet each one of these three is a separate issue that must be addressed before we even attempt to get the other party to buy-in to our change initiative. The second and third questions (agreeing on the solution and the implementation steps) may seem self-explanatory, but it is also vital to make sure that everyone understands and agrees on the problem. What sometimes happens is that in our haste to talk about the change (i.e., the solution), we neglect to verify that we agree on the problem—and if both parties have different problems in mind, the odds are rather slim that our solution addresses their problem. It is no wonder, then, that they fail to see the merit in our solution, and object to it. The three questions of change thus highlight not only what should be covered in a buy-in effort, but also, and just as importantly, the inherent order in which this effort should be executed. There is no sense in talking about the solution before we agree on the problem, and no sense talking about the implementation steps before we

agree on the solution. Hence, the three questions of change act as the basic Layers of Resistance to change that must be overcome or “peeled away,” one after the other, in order to get a buy-in. We use the terms “layers” and “peeled away” since it’s easy to picture the various challenges that must be overcome as peeling away layers of an onion until we get to the heart of the matter: the buy-in (see [Fig. 20-1](#)).

Awareness of the three basic Layers of Resistance is sufficient to improve many discussions about change. It was clear that those guys at the airport were all over the place. The guy who objected kept bouncing from the reasons why it was impossible to implement the change (disagreement on the implementation), to questioning whether they should focus on that particular change (disagreement on the solution), to suggesting that they should solve another problem first (disagreement on the problem). The initiator was doing his best to address each objection, but without any sense of progress; it was no wonder that those two were growing increasingly frustrated with each other and the entire discussion. The first thing they should have done was to pause and make sure that they agreed on what the problem was. Then, once they were on the same page, they could have moved forward to discuss the solution. If at that point they failed to reach an agreement, they would at least know where they stood and could restart from that point. In order to avoid wasting time and trying our own and our partner’s patience, we need to resist the urge to hop all over the place—we should identify as soon as possible the earliest “Layer” on which we disagree, and suggest to the other party that we concentrate on that issue before we move on to the next.

Being aware of where we are in the discussion—identifying the Layer with which we have to deal—may also give us a better idea as to whether we are making progress or we are stuck. In tough situations where the changes may appear “radical” or the other party is exceedingly resistant,

the buy-in process may still take some time. Instead of experiencing the uncomfortable feeling that we are going nowhere, the Layers may serve as a road map, indicating where we are, when it is appropriate to press forward in the discussion, and when we have to take a deep breath and stay put.

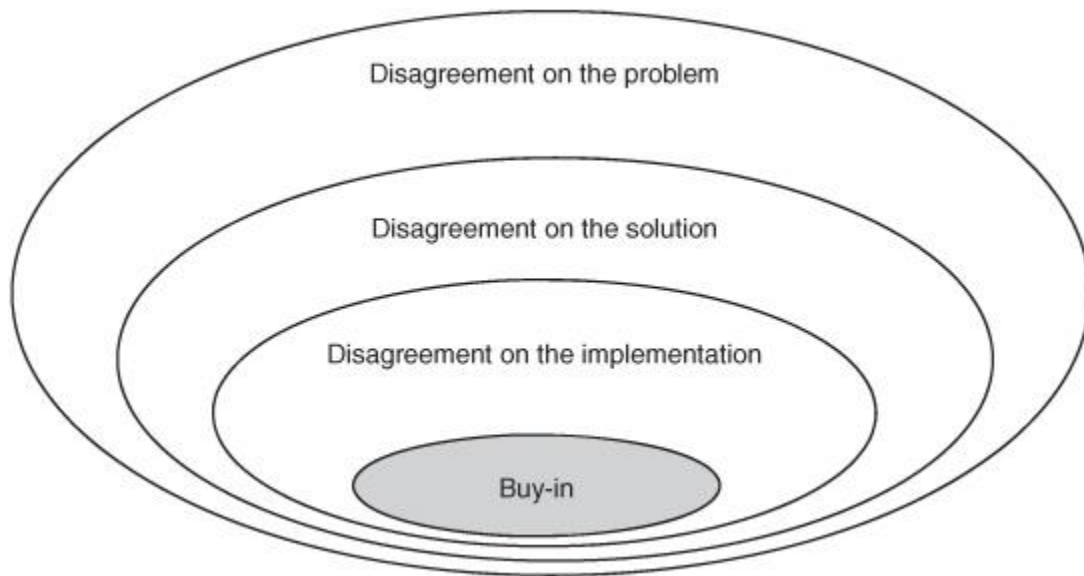


FIGURE 20-1 The basic Layers of Resistance based on the TOC questions of change.

The three basic Layers of Resistance may be the essence of this model, yet they do not tell the whole story. Once we take a closer look at these Layers, we detect even finer Layers inside them. Since the term “Layers of Resistance” was first coined in *My Saga to Improve Production* (Goldratt, 1996a; 1996b, and later reprinted in 2003, 1-14), I have come across outlines of TOC Layers of Resistance that contained within them anywhere between three and nine Layers. The reason for this phenomenon is that in different types of changes, there may actually turn out to be different finer Layers of the basic three that should be dealt with separately. Also, the magnitude of the change has an effect, as large-scale changes tend to have more fine Layers than

local, small changes. Moreover, even with regard to a specific change it is difficult to predict how many and which Layers we will encounter. This is mainly because if we succeed in overcoming one Layer, the other party may overcome the next one independently. In order to develop further our intuition around identifying the Layers and successfully coping with them, it might be worthwhile to review the finer Layers one by one.

Disagreement on the Problem

Layer 0. There is no problem

When we approach the other party eager to discuss the win-win change we believe should be implemented, we sometimes receive responses such as, “What is wrong with what we have right now?” or, “There is no problem,” or, “Everything is fine the way it is.” These kinds of responses clearly indicate that there is no point discussing the problem (i.e., Layer 1) yet, as the other party does not yet acknowledge that there is a problem. We have to take a step back and deal first with Layer 0: Convincing the other party that something is wrong with the current state of affairs. In an illustration that has been used in the TOC community for years when discussing the Layers of Resistance ([Fig. 20-2](#)), we approach Wary Will and tell him, “You have got to make the effort to climb that cliff (read: Implement the change) because there is an alligator right behind you!” Wary Will answers, “What are you talking about? I don’t see any alligator¹.” The only way to move past this Layer is to listen very carefully to what the other party is saying—in other words, to understand what is truly behind their claim. Wary Will may claim there is no problem

because the approaching alligator is still too far away for him to notice it, or he can claim that “there is no problem” because he believes that the approaching alligator is friendly and won’t bite. Because these are two different cases, we will have to use very different arguments to convince Will that there is a problem.

People may be stuck in Layer 0 for various reasons. Sometimes it is because they fail to see that there is something wrong in the current situation. Sometimes it’s the opposite: They may have been well aware of the undesirable effects and fought very hard to get rid of them, but have failed so miserably that as far as they are concerned these negative phenomena must be accepted as part of reality. They might even have become so used to living with these negative phenomena that they no longer see them as negative (Goldratt, 2009, 19). Blockage at this layer may even be inherited from a predecessor who fought and failed, so the person to whom we are talking may not be aware that things can be different. It is no wonder, then, that they don’t think the change we are presenting to them should be a priority.

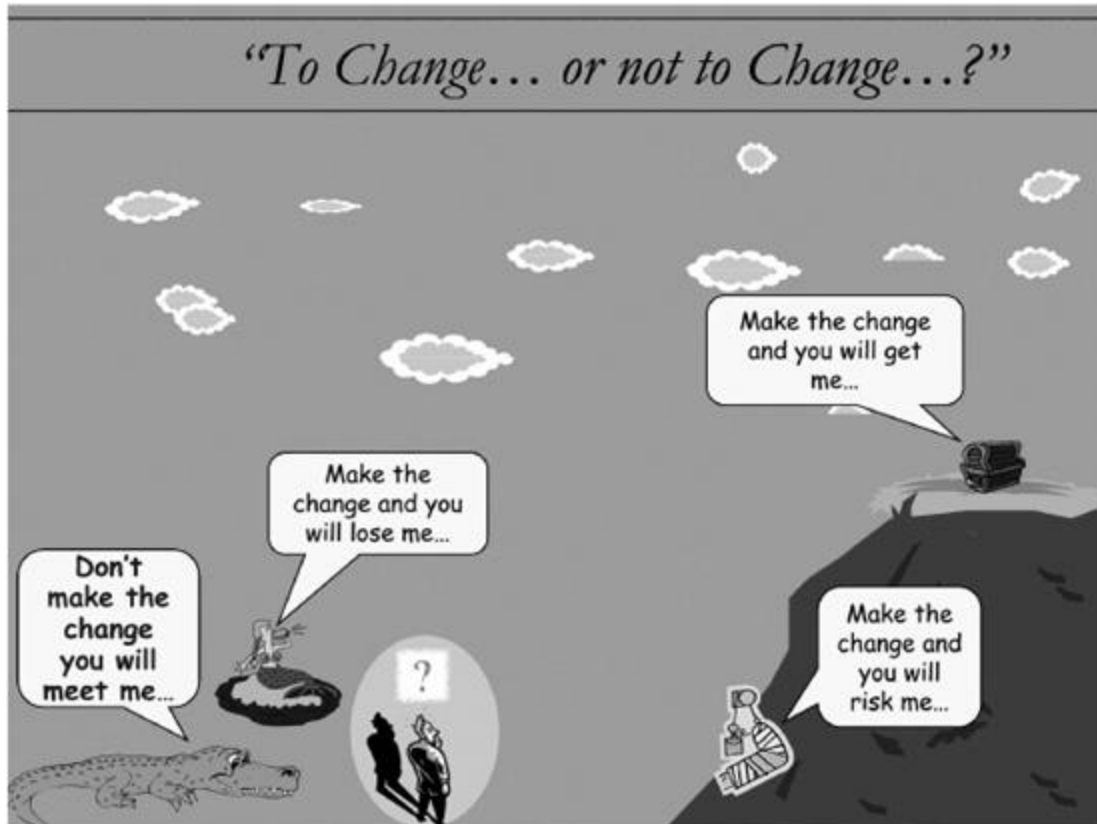


FIGURE 20-2 Weary Will's dilemma: To change or not to change.

How do we move beyond this Layer? The best way is to take the time to understand fully where the other side is coming from. We should let the other party talk and assist them in uncovering their assumptions until we identify their false assumption with regard to the situation. The next step is proving to them that their assumptions are not in fact valid and a problem does exist. Most of the time it only takes a few minutes for the other party to realize that they were operating under a misperception and we can move on. On other occasions, peeling away Layer 0 may take longer. In extreme cases, we might even need to hold a series of discussions in which we gradually bring the other party around to our view of the situation. For those of us who tend to run out of patience, it may be best to muse upon the

alternative for a moment: If we become frustrated, lose our temper, and decide to skip this step, what are the chances of the other person willingly collaborating with us? What are our chances of success without their support?

Sometimes it is wise to prepare some “ammo” going into a discussion with the other party, especially if the buy-in effort takes place as a formal presentation to a group. One approach that might help is to remind the other party of the goal they are trying to achieve and examine whether this goal is fully met with their current mode of operation. If we succeed in making the other party realize that their goals are not met to the extent they would like them to, it means there is a problem—we overcame Layer 0. In other cases, we might consider using another approach to peel back this layer: We need to remind the other party of some significant undesirable effects that are caused both by the problem we are attempting to solve, and by the problem from which the other party suffers. Then we need to convince the other party of two things: First, that these effects do indeed exist, and second, that they are harmful (and thus undesirable). This discussion is not as painful to conduct as it may sound. That is, if we prepare for it. It is best to come up with four to seven undesirable effects that we must verbalize from the other party’s point of view. Remember, mirroring the other party’s terminology is key to getting buy-in. In order to demonstrate that these effects are part of our reality, we can use leading questions, numbers, or any other kind of “proof.” Most of the time this demonstration is sufficient because the undesirability of these effects speaks for itself. And if on occasion one or more of the undesirable effects are not intuitively perceived as negative, we can try to lead the other party through an “if...then” discussion until they realize that those effects are in fact negative.

Now, a word of caution: Let us think for a minute about those people we want on board. We probably need their permission or collaboration because they have some

authority or responsibility in an area that is closely related to our change. It stands to reason, then, that if they have responsibility in an area related to the change, then they are also at least partially responsible for the problem we are trying to solve. So, it might be that they are well aware that there is a problem in the current situation, but they refuse to acknowledge it in public because they don't want to be blamed for it. If this is the case, discussing different undesirable effects and demonstrating how harmful they are might give them the impression that we are blaming them for even more than they thought. It will be like pouring gas on a fire, causing them to resist us even more! That is why we have to be extra careful in the way we approach the other party and in our choice of words. How do we know whether they are ignoring the problem because they are unaware of it or because they don't want to be blamed for it? If we listen well enough, we should know. Buy-in is as much about listening as it is about talking. But what if we are not sure? It is better in this case to play it safe and make very clear to the other side that no one is assigning blame, and all we want is to make things better for everyone.

Let's assume we have peeled away Layer 0 and have gotten the other party to acknowledge that there is a problem. Where do we go from here? Again, we have to listen. If we hear something like, "I see we have a problem, but what exactly is it?" or "Now that I think about it, the problem is different than what you are telling me," it means that we have moved to Layer 1 and should begin to discuss and get agreement on the problem. Sometimes we may find that, especially in small changes, once the other party realizes there is a problem, they are able simultaneously to recognize exactly what that problem is. In this case, we don't need to expend our energy and bore the other side with excessive explanations about the problem. It is better to verify that we are talking about the same problem, realize that Layer 1 is also peeled away, and move on.

Layer 1. Disagreeing on the problem

People come from different backgrounds, have different roles, and have different agendas. Therefore, it is reasonable to expect different answers to the question of what should be improved in a given situation. As we mentioned earlier, it is rather difficult to reach an agreement on a solution unless the two (or more) parties agree on the problem first. If we approach Wary Will and say, "Watch out! There's an alligator behind you!" and he replies, "That's not an alligator, it's a vulture!" then what chance do we have to convince him that climbing up the cliff is a good idea? I have heard people say that to avoid wasting time, it is better to discuss the solution right away and go back to Layer 1 only if during the discussion of the solution we realize that there is a discrepancy in our perceptions of the problem. This shortcut is risky because once we place our cards on the table and the other party objects to our solution, it will be harder to get them to admit they were wrong about the problem in the first place. It is therefore better to play our cards close to the chest to avoid giving the other party the opportunity to object until we verify that we are both on the same page, as far as the problem is concerned.

So, how can we agree on the problem? One way to go about it is to discuss openly the party's assumptions of what the problem is. During such a discussion, we may realize that although we are tackling different problems, they are actually related. It may be that we are talking about the same problem using different terms, or that we are talking about a series of linked problems that should be addressed sequentially. Examining each party's perceptions of the problem enables us to reach agreement on what should be addressed, at which time we can move on from this Layer. Sometimes, if we cannot reconcile the different points of

view, we may resort to negotiating whose problem will be dealt with first. It might work, but it also might result in a stalemate. We may be able to preempt this situation by preparing for it.

Having different roles might mean that different people suffer from different undesirable effects that they mistakenly view as the main problem in the current situation. If we do not deal with the real core problem—the problem that is causing the various undesirable effects—we cannot fully remove those undesirable effects. That is why each TOC analysis begins with a search for the core problem that is causing the undesirable effects in the situation. The TOC thinking tools that are designed to help uncover the core problem are the three-cloud approach and the *Current Reality Tree* (CRT). The buy-in effort can also benefit from this type of analysis. If we can show the other party that their problems, as well as ours, are all derivatives of the same core problem, we may be able to reconcile our different points of view and get consensus on focusing our efforts on the core problem. Whether it is conveyed in a formal presentation or a systematic conversation, getting people to realize what the core problem is and how it relates to their own undesirable effects is very effective in peeling away this Layer of Resistance and enabling us to move forward.

Here again we must be cautious about blaming. We already mentioned that the people we want on board might be sensitive to the subject we are raising. When we approach them to talk about the problem, we might inadvertently give the impression that we are blaming them for the problem. This can easily become the case if we just have an intuitive discussion about it and are not careful about what we say and how we say it. Now imagine what might happen if we approach them with a well prepared, logical analysis that shows how they are not only responsible for their own undesirable effects but are also

responsible for the core problem that is causing everyone else grief and agony. We might be—unintentionally, of course—forcing them into a corner, making them feel blamed or even attacked. And if they get defensive, what are the odds of all of us sailing smoothly toward a happy buy-in for our proposed change? Whether or not they should be blamed is irrelevant at this point. If we are serious about implementing the change, we have to put the issue of blame behind us. Instead, we must concentrate on how the other party is going to perceive our motives for approaching them. They must not feel blamed. Ideally, we want to put them at ease so that they are receptive and positive about our initiative. Being careful about the words we use is the key! If we could also demonstrate that we understand what kept them from solving this problem before, all the better. TOC recommends verbalizing the problem in a conflict format (i.e., a cloud). We want to show the other side not only that we are not here to blame them, but also that we actually understand the conflict in which they are trapped.

When we sense that all sides are on the same page as far as what the problem is, it usually means it's time to move on to discuss the solution. Except, however, for rare cases in which we bump against Layer 2.

Layer 2. The problem is out of my control

Thank goodness this Layer is rare because when it occurs it is very hard to overcome. Layer 2 describes those cases in which the other side insists that the problem is beyond their control and expects us to drop the whole thing. Wary Will tells us firmly, “My hands are tied. There is nothing I can do to help you,” and refuses to hear another word on the matter. When we encounter responses such as these, we had better listen carefully to what the other party has to say. Sometimes the other party is right and the problem is

indeed beyond their span of authority. In order to solve the problem, therefore, we may have to speak to whoever has the power to solve that problem. However, we do not always have the option of approaching someone's superiors, which means we might be stuck.

What if they just say it and the problem is in fact under their control? Here we have a serious problem because they usually refuse to continue the discussion. But if we find a way to open a dialog, we can try to uncover their erroneous assumptions and get them to see the problem is solvable within the boundaries of their control. Or, we can try to convince them, despite their unwillingness, to listen to our solution and then reconsider whether they have the power to implement it.

Disagreement on the Solution

Layer 3. Disagreeing on the direction for the solution

There is often more than one way, more than one "direction" to solve the same problem. Wary Will will probably not help us climb the cliff if he prefers to stay and fight the alligator. Once we agree on the problem, we often bump into Layer 3 (see [Fig. 20-3](#)). What happens in Layer 3 is that each party tries to convince everyone else to go their way. Each party insists that their direction for the solution is better than everyone else's and stubbornly refuses to hear anyone else out. If no one agrees on the direction, there is no point in detailing any of them.

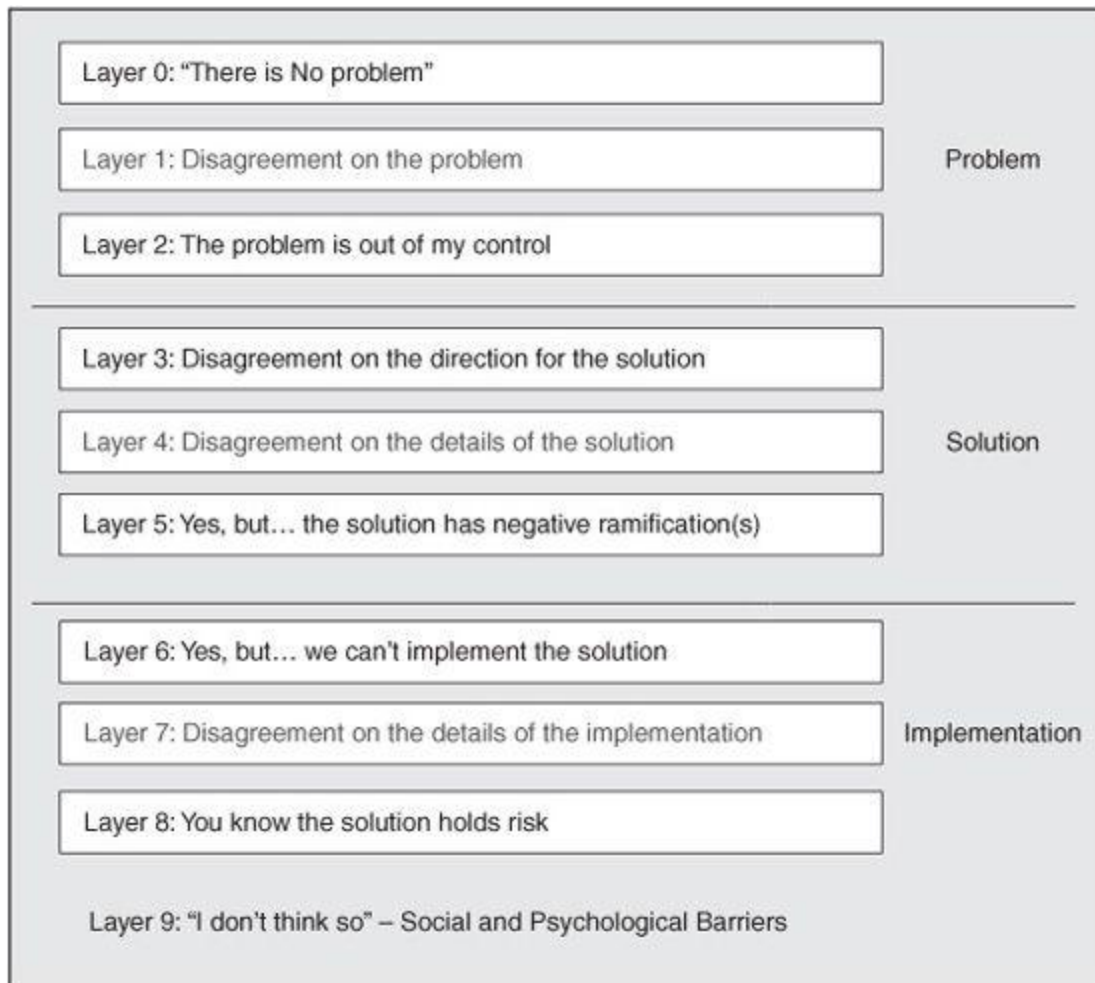


FIGURE 20-3 The TOC layers of resistance to change.

If we anticipate such trouble, we had better come prepared. We need to invest in putting together a list of criteria for what would be considered a good solution. This list may include items such as achieving the opposite of some of the main undesirable effects, meeting the important needs of the involved parties, and avoiding significant negative ramifications. After we present the criteria and agree on them, we should review the directions for the solutions that people have put forth. Since we have invested in identifying the core problem and devised a good solution for it (and we are the ones who wrote the list of good criteria), we have a far better chance of meeting the

criteria than our counterparts. And what if we haven't? Well, perhaps we should realize that their direction for the solution is better than ours is, and proceed accordingly.

It might seem like putting together a list of criteria for a good solution is just a hassle. Why not simply discuss each solution to judge its merit? Many times putting this list together is indeed an “overachievement,” where an intuitive discussion would suffice. But sometimes taking this extra step can make or break our problem solving. It is easy to imagine scenarios involving, well, human nature: If we start comparing solutions, the discussion bears the risk of becoming personal (and emotional) fairly quickly (“mine is bigger than yours!”—sound familiar?). The more we compare and judge, the harder each participant will hold on to and fight for their solution, which makes it much harder to maintain a civil discussion, let alone reach a consensus. A list of good criteria upon which everyone agrees in advance before reviewing any of the solutions serves as a logical fencepost to which we can all refer back. Looking at each solution alongside the list of criteria will help us conduct a practical, rather than personal, discussion. This way we hope that we can let go of the directions that are less desirable and get consensus on one direction. Once we are in agreement on which direction we should take, Layer 3 is peeled away and we can move on.

Layer 4. Disagreeing on the details of the solution

It's important to peel away Layer 3 (the direction for solution) and Layer 4 (the details of the solution) separately when we are facing a change on a large scale that probably has more than one direction for solution and in which there are many details involved in each direction. With smaller, simpler changes, the direction and the details tend to merge

into one discussion about the solution, and trying to keep them separate becomes superficial.

In this Layer we may hear people say, “Your solution is not good enough,” “It does not address the entire problem,” “This is a terrible solution! It doesn’t cover x, y, or z.” People agree to our direction for a solution, but claim the solution is not yet complete; it does not achieve all the desired results. Instead of spelling the doom of our project, such objections actually enable us to check whether we have constructed a comprehensive solution to the problem or we have missed something. We should swallow our resentment toward the other party for poking holes in our precious solution, and instead evaluate their comments as openly (and neutrally) as possible. If their concern is not valid, we should further explain our solution until they see that it is designed to achieve the benefit they pointed out. And if they were right, we should thank them for opening our eyes at this early stage and alter our solution in accordance with their suggestions.

What if we fail to resolve the other party’s issues? If it seems that our plan will fail to achieve a significant benefit, we have to be open enough to re-evaluate our solution and see if it is as good as we thought it was. Maybe we should go back to Layer 3 and choose a different direction for the solution. If we want the other side to evaluate the merit of our suggested change objectively, then we, too, must be objective about it, not blinded by our enthusiasm or sense of ownership.

The other party may bring up more than one desired outcome they suspect is missing. If we are determined to get the other party’s full collaboration (and get the most out of the change), we should listen to what they have to say, determine which significant benefits are missing, and discuss how to modify the change in order to achieve these outcomes as well. One way to systematically overcome this Layer is by first getting consensus on all the benefits (or the

“desired effects”) that the change should bring. To do this, simply write down the opposite of each undesirable effect that was brought up during the discussion of the problem. Then review each desirable effect to determine whether the change is designed to attain it. If one or more of these significant benefits was indeed neglected, we should alter the change to address it. The TOC thinking tool that may assist us here is the *Future Reality Tree* (FRT).

Layer 5. “Yes, but...” The solution has negative ramifications

Once we agree on the solution and believe we have covered all its angles, we are eager to start talking about the implementation steps. This is why we have to take a deep breath when we hear the next expected response—the “yes, but” concerns. “Yes it all sounds good, but you do realize that if we go ahead with this we will end up suffering from...,” they say. I have yet to see one buy-in effort where the initiator did not spend a considerable amount of energy dealing with this Layer. If the other party feels our solution might cause damage, there is little chance they will be willing to collaborate. We must take the time to understand what their concern is and why they claim it is an unavoidable result of our suggested solution. If their concern holds water, we had better address it, and if it does not we should clarify why that is. The TOC tool that is designed to help at this stage is the *Negative Branch* (NBR).

The other party may bring up more than one negative ramification they suspect the change will have. The bigger the change and the more people involved, the more “vulnerable” we are to objections at this Layer. In our haste to complete the buy-in effort, we might look into and address one concern, assume that one small adjustment is enough to overcome this Layer, and move on. This is a grave mistake. If we have not addressed every objection

raised at this stage, the solution will seem harmful and woefully inadequate to the task of solving the problem. Needless to say, the other party will not buy it. There is no getting around it: We must spend as much time and effort as it takes on this Layer until everyone agrees that the solution does not have any significant negative ramifications.

Speaking of negative ramifications, the other party may bring up another type of “Yes, but...” at this point. In this scenario, they may claim that implementing our solution will require them to give up something positive that they already have. Wary Will may realize that by joining us in climbing the cliff, he will have to leave his beloved mermaid behind. No one said a win-win solution was perfect. Sometimes in order to gain new benefits we need to give up ones we have previously enjoyed. At this point in the process, we probably will have already resolved this issue with ourselves and decided that the advantages of the solution are worth giving up some positives. But we cannot make that decision for the other party. Thus, if we truly need them on our side we must convince the other party that the advantages of our solution are worth the price they will pay.

At this point, we already agreed with the other party on what the problem is, and we agreed that the solution we proposed is a good solution. According to TOC, a *good solution* is defined as one that adequately solves the problem without creating new significant problems. In Layers 3 and 4, we verified that our solution will properly address the problem and Layer 5 took care of the negative ramifications. Only now does it make sense to move forward to discuss the implementation.

Disagreement on the Implementation

Layer 6: Yes, but... we can't implement the solution

“Yes, but you’ll never make it,” “It is all fine and dandy but impossible to implement,” “It’s a terrible solution, you’ll never get past x, y, or z.” At first it is difficult to tell Layer 6 from Layer 5 because they both sound the same. However, the objections in those two Layers are very different from one another. In Layer 5, we have not yet agreed that the solution about which we are talking is a good solution. We are still debating about whether it has negative ramifications. In Layer 6, we have already agreed that this is a good solution and we are contemplating how to implement it. People tend to confuse these two Layers more than they confuse any of the others, which results in an ineffective bouncing between objections and a frustrating delay in the buy-in effort. The logical order in which to address these two layers is clear: There is no sense discussing obstacles in the implementation before we agree that this is a change we wish to implement. So, once we go into the “Yes, but...” phase, we need to tune our ears to identify to which Layer the objection belongs, agree with the other party to first address all the negative ramifications, and only then talk about obstacles to the implementation. The way to distinguish between the two types of “Yes, but...” is to ask ourselves, “Is this something that might happen if we implement the change?” (negative ramification), or “Is this blocking me from achieving the change?” (obstacle).

Needless to say, if the other party doesn’t believe that our solution is practical, then there is little chance that they will give us their blessing, so we have no choice but to address all of the obstacles they bring up. As in Layer 5, we have the option between cursing them silently for being a pain or thanking them for making us plan better and face less unpleasant surprises once we go into action. Usually the

bigger the change, the more obstacles we face. And once the obstacles start to mount, we need to sort them out—which obstacles can be tackled in parallel and which have to be dealt with in sequence. The TOC tools that might help at this stage are the *Prerequisite Tree* (PRT) or, in large projects, the *Strategy and Tactic Tree* (S&T).

Layer 7: Disagreement on the details of the implementation

As in the case of Layers 3 (direction for solution) and 4 (details of the solution), Layers 6 (obstacles to the implementation) and 7 (details of implementation) should be addressed separately when planning large-scale changes. In small changes, they tend to merge into one Layer that covers our attempt to reach an agreement on the implementation plan. At Layer 7, we discuss and get consensus on the little details: schedules, due dates, assigning roles and responsibilities, budget, resources, etc.

Deciding “who does what” is something we all do fairly well. However, we should not neglect the “why.” Explaining the logic behind our decisions is not only helpful in convincing people that our plans make sense, it also facilitates high performance. We can be as nitty-gritty as we possibly can, detailing exactly what to do where and when, but reality may not turn out the way we expect it to and these details might be worthless. Change holds considerable uncertainty and the effective way to handle it is not by presenting tiny specifications but by providing the “why.” If people understand why we want them to do something, what each step is aimed to achieve, and why they need to do it before moving to the next step, they will be in a much better position to improvise successfully when reality doesn’t turn out the way we expected it to. The TOC tool that may be helpful in conveying the “why” of the various Implementation steps is the *Transition Tree* (TRT).

Delegating tasks in this way tends to motivate people, which also has a positive impact on their willingness to collaborate.

Layer 8: You know the solution holds risk

As we go through Layers 6 (obstacles to the implementation) and 7 (details of the implementation), the other party may become aware of possible risks that we take if we decide to go ahead with the change. Wary Will realizes that we want him to climb up a shaky ladder and he immediately responds, "I don't know about that, I might break a leg." As long as the other party believes the risk is not worth it, we are in trouble. It is up to us to discuss each risk they bring up and think how we can lower it by making some changes (e.g., fixing the ladder for Will) or creating safety nets (e.g., placing a mattress below the ladder). If we can't find a way to lower the risk, we need to reconsider the way we decided to implement the change (maybe there is an available zeppelin in the area?). If we can't find a way around that risk, we might end up in a position where we need to weigh the risk against the potential damage of canceling the plans for implementing our solution and see what will be the best course of action. Needless to say, if we want the other person's collaboration, we had better convince them that we have made the right decision.

Going through this buy-in process significantly improves the odds of convincing the other party to go along with us. The logical order and the intuition to know where to pause and how to handle each type of objection provides us with a way to better master this dialogue. Utilizing the Layers of Resistance also gives us much more control than we would have if we conducted such discussions in the intuitive way. In the intuitive way, after we present the change we usually resort to addressing whatever objections the other side

raises, so we are actually giving them control over the discussion. Utilizing the Layers allows us to know which layer we are in and what we should talk about, so that when they raise another objection we can tell if it belongs to an earlier Layer or to a later one. We then know if we need to go back, or we need to show the other party we heard them (preferably write their objection down), and explain why it makes sense to postpone dealing with it until a later stage. This way we remain in a better position to steer the conversation.

And what if we got this far, covered all eight Layers, and the other party still resists? Here again we have to listen very carefully to what they say. The first thing we have to consider is that we may have lost them at an earlier stage and they are still stuck there. If this is the case, evidently we need to go back and pick up the ball from where we dropped it. Another cause for resistance at this point is that they simply need time. Often people are not comfortable with giving their blessing right away. They need to take their time to think it over and after they get used to the idea they will most probably come back to us with a positive answer. If it is not a problem at an earlier layer and it's not the need to digest, resistance at this point means we clashed against Layer 9.

Layer 9: "I don't think so"—Social and psychological barriers

The Layers of Resistance provide order to the objections that relate directly to the change at hand (i.e., inherent objections). However, we cannot ignore the fact that people may also resist due to reasons that are not inherent to our change (i.e., external reasons). As was mentioned at the beginning of the chapter, people may possess personality traits that make them more prone to resist change. People may feel pushed out of their comfort zone and resist the

excessive (perceived) uncertainty. People may resist because of social pressure or because they conform to social norms that our solution challenges, or because of various other reasons. Whatever the external reason is, it may stand in our way from the very beginning of the buy-in process, but as long as we haven't addressed the inherent objections, we should not focus on it (see again [Fig. 20-3](#)). As tempting as it may be to cling to it, all it does is lead us to blame the other party instead of take the responsibility to buy them in.

Think, for example, of a case where we would need to present an innovative change that contradicts the way things have always been done. Let's assume that we detect fairly quickly that the other party objects because they prefer to stick to tradition and conform to the way others in the field behave. One way to react is to call them "conservative" or even "primitive" and . . . and then what? Another way to handle the situation is to have faith that this external reason might not block us but merely slow us down. We should acknowledge it might take more effort, but nevertheless attempt to buy them in. If conservatism is their only reason for resisting the change, often we find that these people eventually come around; they realize there is a problem in the current state of affairs and even if our solution contradicts the traditional way of doing things, it is in fact the right thing to do.

The earlier we detect an external reason for resisting, the better we can fine-tune our approach in order to overcome it. We use the Layers of Resistance while keeping the external reason in mind. For example, if we realize we are pushing people out of their comfort zone we should continually ask them what information they are lacking, and discuss how to make things easier for them in the implementation stage by using demonstrations, pilots, etc. Or, if we realize we are talking to a person who needs to be in control (and we can't go around them or shoot them), we

have to alter our approach to give them more control—both in the buy-in discussion and in the implementation of the change.

When we bump into Layer 9, it means that we have done our best to take the other party through the Layers that deal with objections inherent to our change, and we are now convinced that the reason they still resist is external to our change. In this situation, we should identify the external reason for resistance if we haven't done so earlier, and attempt to address it. The purpose of this chapter, however, is not to cover a comprehensive list of external causes for resistance to change, as there is plenty of literature on the subject.

Sense of Ownership: The Key to True Buy-In

There is one type of change people are truly excited about—their own initiatives. As we are well aware, psychological ownership (“this is MINE!”) plays a key role in people’s enthusiasm and commitment. Thus, the more important the change is to us and the more collaboration we need from the other party, the more we should invest in making them feel this is “their” change too. The problem is that when we initially ask for their collaboration, they have no sense of ownership; they feel as if they have nothing to do with this change. How do we cultivate this feeling? A sense of ownership may emerge through various related routes (see, for example, Pierce et al., 2001²).

Using the Layers of Resistance can be an excellent way to build a sense of ownership; that is, if we are truly willing to share the ownership of our change with others. The way to go about this is to set aside our egos and learn to welcome inquiries and objections. After we present our ideas in each Layer, we need to encourage the other party to ask questions. This is *not* about asking questions for the sake of

asking questions. This is about encouraging the other party to speak their mind so we know what is truly bothering them. Discussing what is bothering them and clarifying the missing details is what will help them become familiar with the change. In addition—and this is the key to the whole thing—we have to evaluate their objections objectively. Keeping an open mind, we will find that at least some of their concerns hold water. If we accept their reservations and ask for *their* input of how to overcome them, we give them control over current decisions and future actions. The more we acknowledge their (valid!) reservations and incorporate their suggestions into the change plans, the more it will become their change too. Even if we have effectively identified the problem and come up with a reasonable solution, the other party will most probably raise valid concerns in Layer 5 and real obstacles in Layer 6. Instead of trying to dispute these concerns, we should view them as excellent opportunities for building the other party's sense of ownership. When such a discussion is done well, the other party feels more involved and more willing to participate once we get to Layer 7. If at that point they assume responsibility and start taking charge in reviewing the little details, we know we have made it.

Another issue that should be discussed here is the issue of fairness. The other party may have resisted our change all along because they have been trying to get more out of it for themselves. They might believe they deserve more because they have invested a lot in the past and feel they weren't adequately compensated, suspect they will be asked to invest a lot in implementing the change, or might believe that others receive more. The issue of fair compensation for their efforts or fair distribution of the expected outcome from the change might be especially complicated if we are dealing with a group of people—the ones who expect to invest more in the implementation of the change and claim they deserve more based on their

contribution, the ones in the middle who will advocate the equality principle, and the weak ones who expect to contribute the least and will claim they deserve more because of their special needs. The issue of what is fair and not fair is a muddy swamp. If we go there, we are much more likely to drown than float. What might help is to build peoples' sense of ownership in the change. If *they* offer to help or *they* decide they would like to invest more, the issue of fairness may not come up.

Apart from promoting a sense of ownership, there is another big advantage to welcoming objections and evaluating them objectively. We have all implemented changes we were excited about just to find out later that they were “half baked” and did not yield all the desired results. If we truly listen, there is a good chance that other people might be on to something that we have missed—thus, they increase our chances of implementing a well-planned change and fully enjoying its results.

Bottom Line

Reading about this buy-in process might give the impression that persuading people is a complicated task that takes a lot of work. Well, sometimes it does, but let's put things in perspective. Most everyday changes are local, small changes that require no more than a good open discussion. In this type of change, we usually encounter no more than three or four Layers. By being aware of them, the discussion tends to be more focused and the buy-in effort actually takes less work.

When we face a large-scale change, things are different. Here, we might need to invest a significant amount of time in preparing our presentation and planning how to conduct the buy-in discussion. When faced with these preparations, we might think that it is too much effort and decide to “wing

it.” The perception of this effort being “too much” comes from comparing the time we need to prepare to the time it will take to “wing it.” But if we look at the big picture, what we should compare is the time it will take to prepare our analysis and discuss our arguments to the time, effort, and agony we will go through convincing the other side to say yes (and we may never even hear that blessed word!) if we don’t prepare. Try to recall a time when you have encountered resistance and how hard you worked to get the other party to collaborate. If only you had done a little more homework before you had leapt in

References

- Goldratt, E. M. 1984. *The Goal*. Great Barrington, MA: North River Press.
- Goldratt, E. M. 1996a. “My saga to improve production: Part 1,” *APICS—The Performance Advantage*. 6(7) (July): 32-35.
- Goldratt, E. M. 1996b. “My saga to improve production: Part 2,” *APICS—The Performance Advantage*. 6(8) (August): 34-37.
- Goldratt, E. M. 2003. “My saga to improve production.” In *Production: The TOC Way*. Revised Edition. Great Barrington, MA: North River Press.
- Goldratt, E. M. 2009. *The Choice*. Great Barrington, MA: North River Press.
- Pierce, J. L., Kostova, T., and Dirks, K. T. 2001. “Toward a theory of psychological ownership in organizations,”

The Academy of Management Review, 26(2)(April):
298-310.

About the Author

Dr. Efrat Goldratt is an organizational psychologist who specializes in the *Thinking Processes* (TP) according to TOC. She has played an active role in the development of the TP, especially in applications for individuals. She has been teaching the TP for both business and education worldwide.

Dr. Goldratt has a PhD in organizational psychology and conducted her dissertation research on employees' reactions to positive organizational change.

CHAPTER 21

Less Is More—Applying the Flow Concepts to Sales

Mauricio Herman and Rami Goldratt

Introduction

Ever since we embarked on the Viable Vision, the mission of our company is to become an ever-flourishing company. A company that is capable of exponential growth by relying on processes that ensure growth does not come at the expense of stability. Our main drive is to establish, capitalize, and sustain a decisive competitive edge. A decisive competitive edge can only be created by satisfying a significant need of the market to the extent that no other significant competitor can. The needs we chose to capitalize on are Reliability and Speed.

During the last years, we have fully transformed our sales approach. We have changed from selling products to selling solutions, from selling based on price to selling based on value. Our sales people understand more and more that chasing orders is not the key for growth; that they should possess the skills to close business deals. We have implemented changes not only in the way we are selling, but also in the way we administrate and manage sales opportunities.

The results of these efforts can be viewed in many aspects. Our sales and Throughput have increased in recent years. The market (based on the reaction of clients)

perceives more and more that our company is not just another supplier. We have significantly changed our product mix to better Throughput products. We have increased our share with existing clients. We have increased our client base and reduced our dependency on a few big clients. The efforts have also produced less tangible results; looking at the behavior of our departments, it is evident that we have dramatically enhanced our ability to initiate and adopt changes.

Despite these results, we all felt that something was still lacking. The most important measurements—profitability and sales volume—were a clear indicator that there was still a big gap between the current reality and the reality we wished to create. When viewing the trend of sales and profit growth, it was apparent that we were not growing at the desired rate. The obvious question was, “What is still missing?” With all the improvements we had made, why were sales not growing at the much faster rate that all indicators showed we should have achieved?

Copyright © 2010 by Mauricio Herman and Rami Goldratt.

In the last few months, we have improved the measurement and reporting of our sales efforts. One striking measurement was our hit ratio. Our funnel seemed loaded with opportunities, but we were winning very few of them. Every client we presented our offer to liked the offer, and eventually we won at least one of the requests for projects¹ that the client requested from us. However, most of the requests by these clients that entered the sales funnel did not turn into orders. Moreover, the same clients who expressed genuine interest in our offer were evidently introducing many promotions to the market without us even participating in the process. The number of projects in our sales funnel was 250; our hit rate was at 11 percent.

In July 2008, I read Eli Goldratt’s new article, “Standing on the Shoulders of Giants.”² The article allegedly deals with

production issues; it highlights the concepts that underlie three of the major production system breakthroughs of our times: Henry Ford's production lines, Taiichi Ohno's *Toyota Production System* (TPS, later known as Lean), and the *Drum-Buffer-Rope* (DBR) application of the *Theory of Constraints* (TOC), which was developed by Eli Goldratt. You may ask yourself why this is relevant to the topic of sales. Well, the concepts underlying these breakthroughs struck me as being highly relevant to the management of sales opportunities.

The following paragraph is taken from the article:³

In summary, both Ford and Ohno followed four concepts (from now on we'll refer to them as the concepts of supply chain):

1. Improving flow (or equivalently lead time) is a primary objective of operations.
2. This primary objective should be translated into a practical mechanism that guides the operation when not to produce (prevents overproduction). Ford used space; Ohno used inventory [Goldratt uses time].
3. Local efficiencies must be abolished.
4. A focusing process to balance flow must be in place. Ford used direct observation. Ohno used the gradual reduction of the number of containers and then gradual reduction of parts per container. [Goldratt uses buffer time consumption]

Improving Flow

Our question was, “Do these concepts apply to the sales funnel management environment?”

The “work” flowing in the sales funnel is sales opportunities. How important is it to ensure opportunities flow with as few disturbances as possible through our process? Just like in production, delays in flow translate to longer lead times. In both environments, longer lead times means poor service to clients; it means deferred income; it means some of the entities flowing (be it work orders or sales opportunities) suddenly become urgent, and so on. In sales, just like in production, delays in flow often entail higher cost (be it work-in-progress [WIP] inventory or sales expenditure).

On top of the implications on cost, it is commonly known that when the system is clogged with WIP it gives rise to quality problems (masking them and making them more difficult to manage). In essence, the same goes for the management of sales opportunities. Delays in flow of opportunities typically entail quality issues, as salespeople and sales support functions need to deal simultaneously with more opportunities that are not flowing smoothly. It is apparent, therefore, that all the reasons why flow is important for production apply to the management of sales opportunities as well.

However, there is one striking difference. As a matter of fact, in sales, flow is of much greater importance. Unlike in production, the longer a sales opportunity is delayed in a certain step, the lower is the probability to win this opportunity. Moreover, when opportunities are not flowing, more time and attention are required by the salesperson or the function dealing with these opportunities. Try to imagine that in production the longer the work order is in the queue, the longer the touch time becomes. In sales, this *is* the reality. The attention given to delayed opportunities is at the expense of bringing in and following up on other opportunities. Flow, therefore, is certainly a primary

objective when it comes to the management of opportunities in the funnel.

What about the second concept of supply chain?

Preventing Overproduction

The second concept (*preventing overproduction*) is known by our production people as “*choking the release*” (*not releasing work to the floor until a certain time—the buffer—before its due date*). The underlying assumption is that having too many orders on the floor creates jams, masks priorities, and disrupts the flow. Is this relevant to the sales environment? Let’s examine the ramifications of having too many open projects. Having many open projects means that every resource involved in the sales process is simultaneously responsible for performing tasks across multiple projects. When a resource is working on many projects, bad multitasking is unavoidable; the resource jumps from one project to another without really advancing any of the projects. When different resources need the inputs of each other to complete their tasks, bad multitasking intensifies. To complete the task, one needs the input of the other (which, for example, can be a designer, a buyer, an account manager, or the client), but since the other is not available (busy on another task), the first resource jumps to another task. When the second resource becomes available, the first one is now busy on the other task, so the second resource jumps again to another task, and so forth. Basically, the two resources are frequently waiting for each other. Bad multitasking significantly increases the cycle time and derails the attention (and with it the quality of work) given to each opportunity in the funnel. When the response time is longer and the quality of work is reduced, the chance to turn opportunities into orders is significantly reduced.⁴

You may think this is not a major problem in handling opportunities in the funnel because one often knows which projects to focus on, and by that can avoid bad multitasking. In fact, our dear salespeople are smart and often have enough experience to tell early on which opportunity is more interesting to the company (meaning it is real, it will be realized in the short term, it yields good Throughput, and it is not a complicated project that would risk our performance). It is not surprising, therefore, that more attention is given to those opportunities that are experiencing a higher hit ratio and a shorter sales cycle. In our company, I'm convinced, many thought that this was actually a proof that we had managed the bad multitasking because it seemed that we were really focusing on the good opportunities. However, on this point we were terribly wrong. We were completely blind to the negative ramifications the immense number of open opportunities had on the attention given to processing good opportunities, and more importantly, on the attention given to introducing more good opportunities.

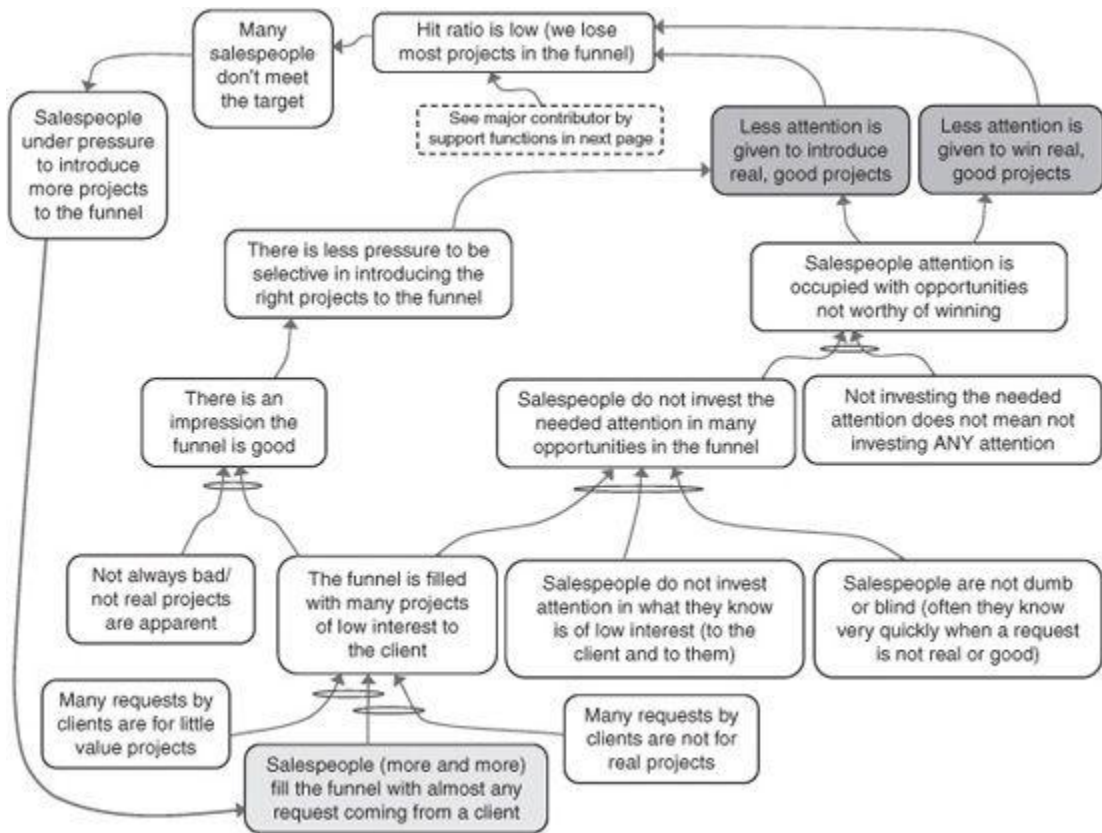


FIGURE 21-1a Introducing almost any request into the sales funnel.

The cause-effect diagram (*Current Reality Tree* [CRT]) in [Fig. 21-1a](#) describes the ramifications of introducing almost any request by the client to the funnel.⁵

[Figure 21-1b](#) shows the ramifications on sales support functions such as engineering.

As we can see, a starting point to this CRT (cause-effect diagram) is the phenomenon, “Salespeople fill the funnel with almost any request coming from a client.” Why did we feel the pressure to do so?

Since I believe people in general (and definitely in our company) work with good intentions, there must be a positive need that drives this behavior. As is logically shown in the CRT, the need that drives us to fill the funnel with almost any request coming from a client is, “Ensure enough

opportunities.” We assume that to ensure high sales volume, we should take advantage of any opportunity that we have and not limit the funnel. Since the hit ratio was low, we believed that we needed to introduce as many opportunities as possible to the funnel in order to reach the target. We did this even when we were skeptical about the validity or value of the opportunity we introduced. We hoped that some of these bad opportunities would turn good. We hoped that the client would eventually give us good opportunities as long as we interact with him, so we accepted the pseudo-orders he requested. We assumed that to reject a request coming from a client would hurt the relationship. These assumptions are not coming from thin air; they are based on anecdotal instances that we encountered in our engagements with clients. Of course, we also assumed that there are not enough good opportunities around to generate the needed volume. To summarize this point, “In order to ensure enough opportunities in the funnel, we believe we must fill the funnel with almost any request by a client” ([Fig. 21-2a](#)). This practice was perceived as a necessary condition in our reality to generate the desired volumes.

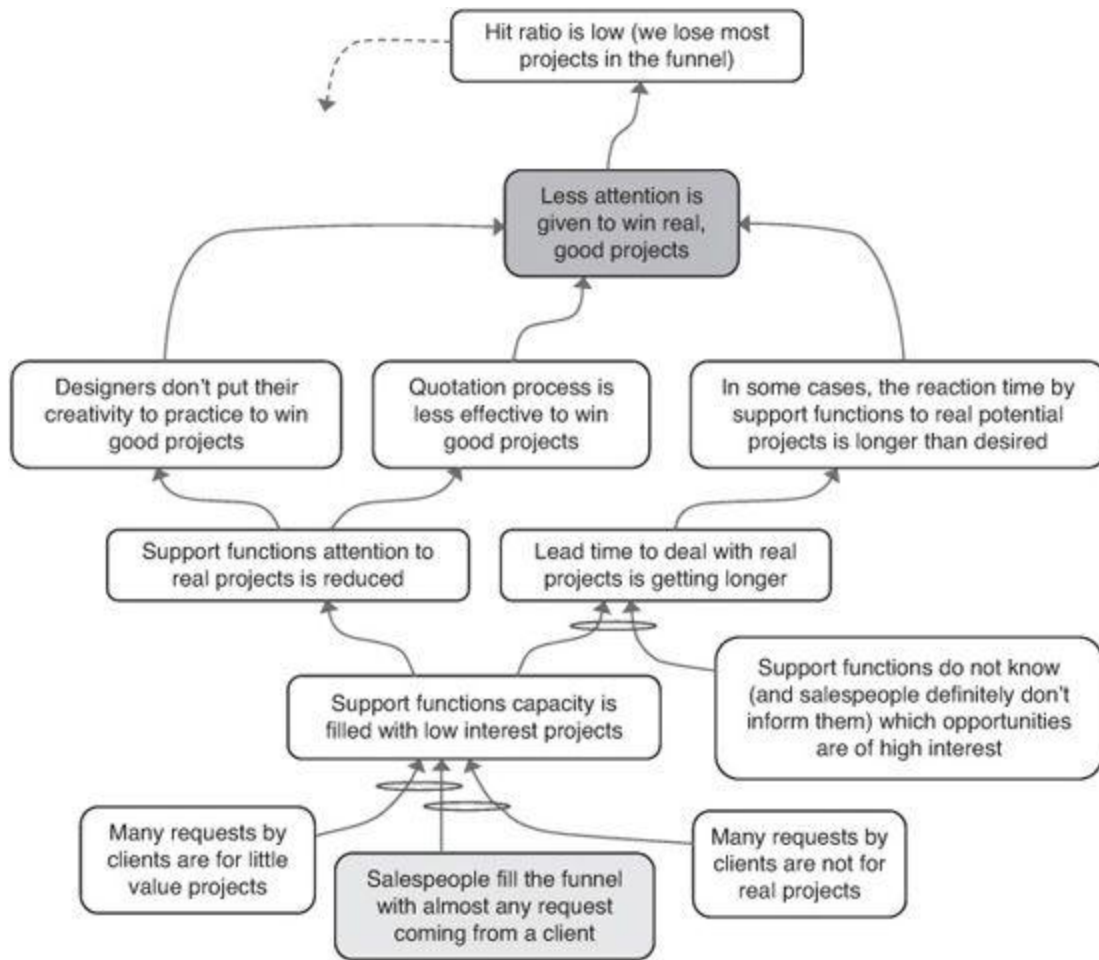


FIGURE 21-1b Effect on sales support functions.

We did not pay attention to the negative ramifications of doing so. The fact that flooding the sales funnel with opportunities leads directly to bad multitasking on both salespeople and support functions created a false impression of the funnel and masked priorities. As explained previously, the impression that we are able to focus only on the good opportunities and by that avoid bad multitasking is an illusion. Flooding the funnel unavoidably leads to less attention being paid to bringing in and following up on real, good projects. Note that not only is our ability to win opportunities jeopardized by bad multitasking, but also there are many good opportunities in the market that require more attention by the salespeople in order to expose

and win them. An example could be a very good project that a client for some reason is contemplating carrying out with another supplier, and therefore we would not hear about it unless we devote time and attention to expose it. Another common example is a project that is being managed by other personnel within the client's organization to whom we are not currently talking. It is highly important, therefore, to notice that bad multitasking on the current opportunities in the funnel also has devastating effects on the ability to introduce more good projects. In essence, if we desire to win more high-value projects (have better projects and increase the flow), we should limit the number of opportunities ([Fig. 21-2b](#)), selecting very carefully to what to devote our attention.

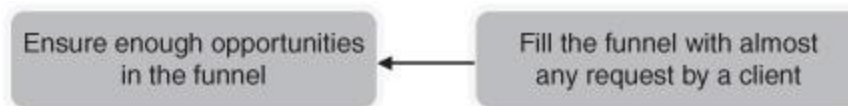


FIGURE 21-2a Fill the funnel.

Now, the conflict is clear, as we see in [Fig. 21-2c](#).

According to the second concept of supply chain, the primary objective of flow should be translated into a practical mechanism that guides Operations when not to produce (prevents overproduction). In our scenario, this means limiting the number of opportunities in the funnel. We now understand that what prevented us from doing so is the above conflict—the fear that limiting the number of opportunities would result in not having enough opportunities in the funnel to generate the desired sales volume.

But is this fear really valid? We assumed that in order to have enough opportunities, we should flood the funnel with opportunities. It is a chicken-and-egg situation. Flooding the funnel results in the low hit rate that leads us to introduce

more and more opportunities. This loop continuously makes us believe that to have enough opportunities to generate the high sales volume we need to have many, many opportunities in the funnel. However, as we see in reality, this never brought us to the targets we have set. Actually, by introducing more and more opportunities, we were not getting enough orders to reach the high sales targets.

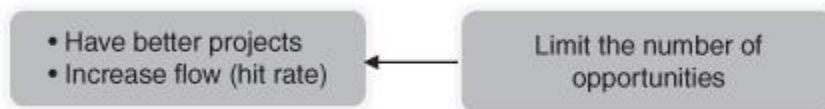


FIGURE 21-2b Limit the number of opportunities.

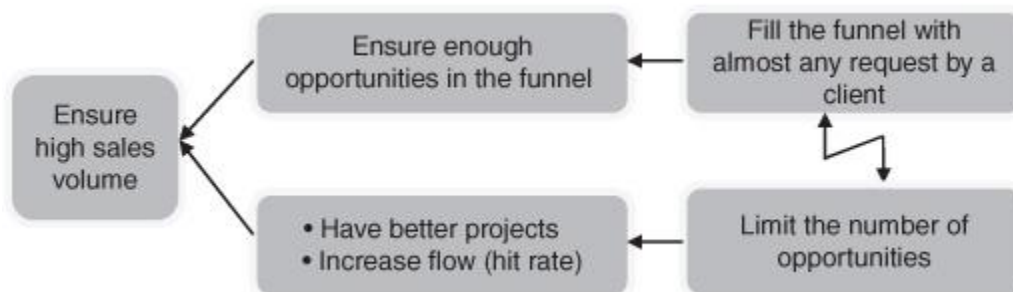


FIGURE 21-2c The dilemma of filling the funnel versus limiting opportunities.

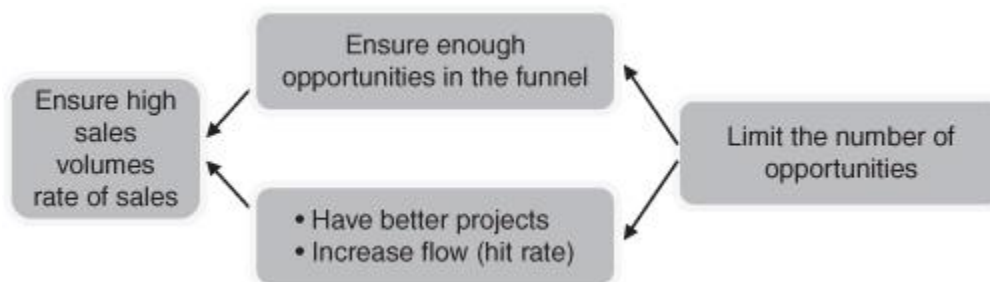


FIGURE 21-3 Limit the number of opportunities.

Think what would be the case if the funnel would be occupied with good opportunities and better attention would

be provided to each. Would we still need to have as many opportunities in the funnel to reach high sales volumes? Limiting the number of opportunities in the funnel would result in providing much better attention to each opportunity and induce us to look for and introduce good projects. If this is the case, then in order to have enough opportunities to reach the sales volume, we don't need to introduce every request for a project we receive. In fact, we should actually limit the opportunities in the funnel, as in [Fig. 21-3](#).

As we just concluded, it makes sense to limit the number of opportunities in the funnel. The question then becomes how to do it. At the early stages of the process (where most bad opportunities lie), we do not have a due date that would determine the release point of the opportunity to the funnel in the same manner in which our production system operates.⁶ We needed a different mechanism to limit the number of opportunities in the funnel. Here we turned to the Project Management solution of TOC.

As you know, bad multitasking is prevalent in multi-project environments, such as R&D or maintenance departments in any company, where shared resources are working on many projects in parallel. The solution to reduce dramatically the bad multitasking in such environments is simply to set a maximum number of open projects (even if it means freezing existing projects). Only when a project is completed is a new project opened. We decided to follow the same approach. We would determine a maximum number of open projects in the funnel. Obviously, this number must be dramatically lower than the number of open projects currently in our funnel; otherwise, we would not reduce bad multitasking. During a meeting we had with all sales directors, we decided to set this limit to 50 percent of the existing opportunities in the funnel. When we set this maximum number, we used our intuition and followed a rule

of thumb. (We also had predicted that it would not be immensely difficult to “freeze” or take out 50 percent of the opportunities, as most of the opportunities are not real or attractive. This prediction was evidently valid, as it took us only an hour to make the decision and determine which projects to remove from the pipeline.)

In retrospect, our intuition was guided by the same logic underlying the “calm-between-the-extremes curve” that exists in multi-project environments.⁷ Choosing to have more opportunities in the funnel elongates the sales cycle and increases WIP, but since more opportunities means more safety buffer to recover for lost opportunities, expectations are that a high number of opportunities will be won. This is correct when not a lot of opportunities enter the system, but when the amount of opportunities is considerable, another phenomenon starts to raise its ugly head. What we have to bear in mind is that the higher the number of opportunities, the lower the attention given to each one. When there are too many opportunities in the funnel, bad multitasking starts to occur. The higher the bad multitasking is, the lower the hit ratio is.

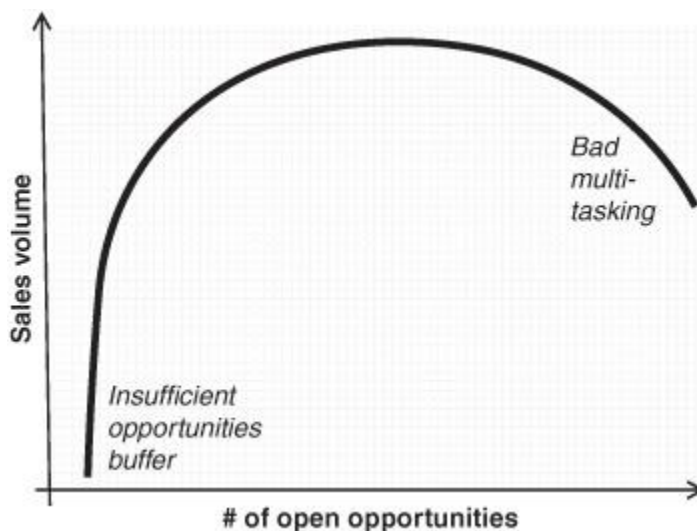


FIGURE 21-4 Calm between the extremes.

The magnitude of generated sales volume as a function of the number of open opportunities is shown schematically by the following calm-between-the-extremes curve in [Fig. 21-4](#).

When one wishes to determine the number of projects to cut, one needs to be very careful not to go overboard. In other words, do not bring the environment from the extreme right side of the curve—where it is—to the extreme left side. The following formula would do the trick:

$$(\text{Number of Opportunities} \times (1 - \text{Hit Ratio}))/2$$

Since being on the extreme right side of the curve assumes high bad multitasking and therefore a very low hit ratio, following the above formula would bring the number of open opportunities to be between the two extremes. If the hit ratio is not as low, the number of projects that would be cut according to the formula is reduced to avoid reaching the left extreme. In our case, since the hit ratio was 11 percent and the number of open projects (opportunities) was 250, if we would have followed this formula we would have cut practically the same number of projects as our intuition guided us to do.⁸

First guideline: Choke (and even freeze/cut) the number of open opportunities each sales division has in the funnel, and set it as the maximum number of opportunities a sales division can hold in the funnel.

What about limiting opportunities on the salesperson level? How are we going to make sure most of the opportunities for a given division do not fall on the shoulders of a few sales-people, leading to bad multitasking? The pragmatic answer for our company was that until we see there is a problem that calls for a policy, each director determines when a salesperson is handling too many opportunities, and then opportunities should be handed over from one salesperson to another. Our sales

measurements and incentives needed to be adjusted to allow this to happen.

Local Efficiencies Must Be Abolished

Let's now turn to the third concept of supply chain: "Local efficiencies must be abolished." First, let's understand it. One of the major enemies of flow is "local efficiencies"—the perception that any point in the chain must work as much as possible. In essence, it corresponds to the erroneous view that encourages measuring the load (number of opportunities) of the funnel instead of measuring the output of the funnel. Examples of local efficiencies could be measurements like:

1. Number of sales calls/opportunities each salesperson has—the more the better.
2. Number of opportunities in the various stages of the funnel—the more the better.
3. Number of projects a designer is working on—the more the better.

We need to make sure we are not using measurements or policies that aim to increase local efficiency and by that jeopardize the flow of opportunities in the funnel.

Second guideline: Stop incentivizing to increase the number of open projects in the funnel. Check if there are other local efficiency policies, measures, or behaviors that jeopardize flow.

We applied the first three concepts of supply chain to our sales funnel management approximately 3 months ago (mid-July). We expected that our hit ratio and sales-cycle duration would be improved as better attention would be given to each opportunity. We speculated that the

Throughput per order would grow as better projects would be introduced. And, of course, we predicted that sales would grow as the flow of better projects would be dramatically improved.

We would like to be very cautious about concluding the results achieved, as they have been way above what we have expected. The following results (also presented in graph form) achieved in the last three months since we have implemented the choking, are measured on a rolling five weeks average:

- Hit ratio increased from 11 to 40 percent ([Fig. 21-5](#)).
- Sales cycle duration shortened from average of 32 days to 17 days ([Fig. 21-6](#)).
- Average Throughput per order grew from 52 to 68 percent ([Fig. 21-7](#)).

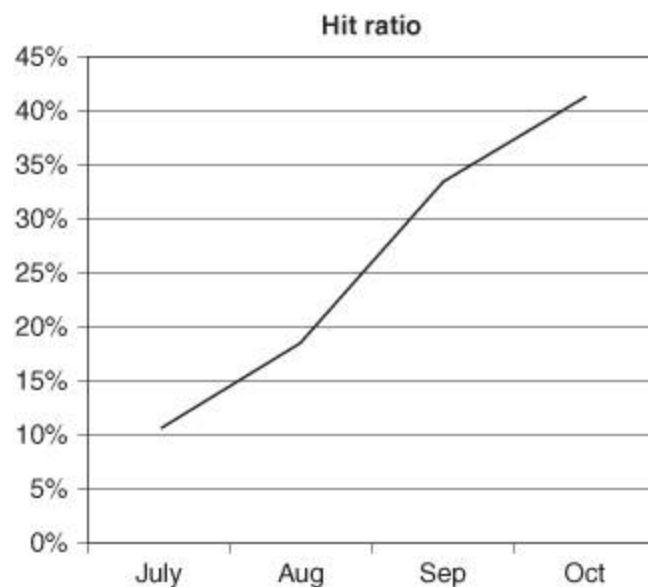


FIGURE 21-5 Hit ratio

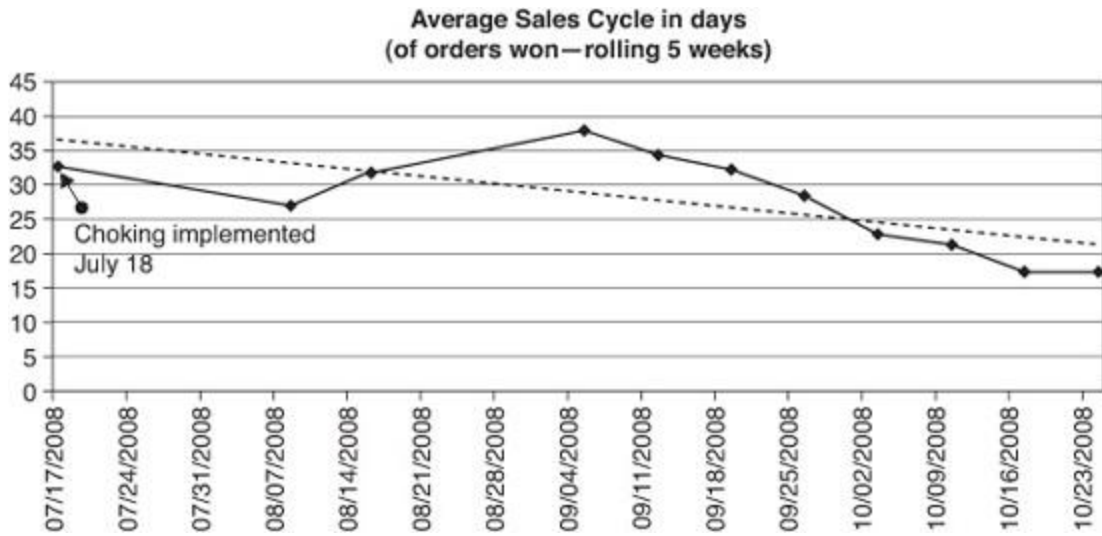


FIGURE 21-6 Average sales cycle in days.

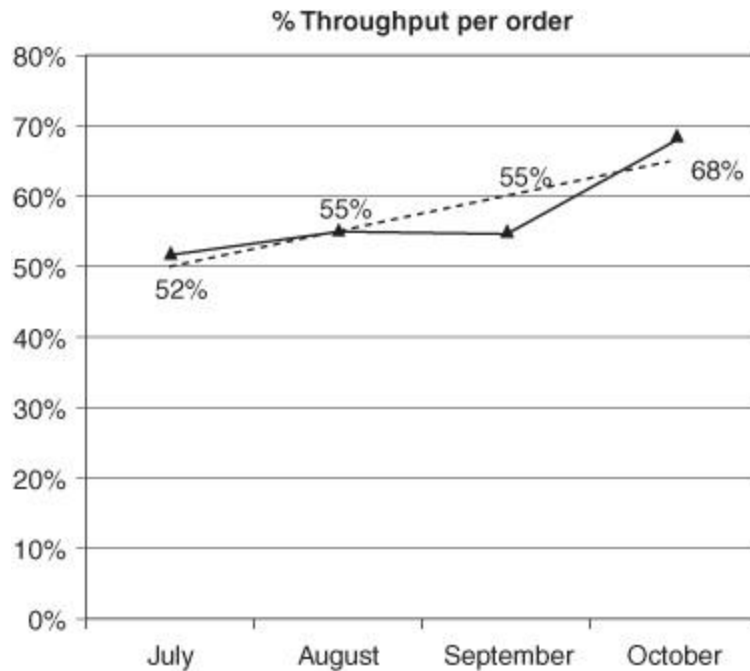


FIGURE 21-7 Percent Throughput per order.

- What about sales? Here we need more time to assess the effect, not because sales have not grown. On the

contrary, we know that sales have grown by much more than 20 percent. However, this growth had its effects on the plant. We have learned the bitter lesson of not contemplating the negative effects of success. In October, we had to postpone many orders to November, we had orders being canceled, and our salespeople's attention was shifted to dealing with not-so-pleased clients—to say the least (I estimate this has occupied at least 30 percent of their time). It will take us two more months to assess the magnitude of the growth in sales.

These results were achieved by applying the first three concepts of supply chain. The following is a description of the way we are going to apply the fourth concept. It should be read, therefore, as a possible way to apply it, and not as a model that has already been tested and proven.

A Focusing Process Must Be in Place

The fourth concept of supply chain states the following: “A focusing process to balance flow must be in place.” In practice, balancing the flow means to eliminate any major disruption to the flow. In production, disruptions become apparent by the accumulation of WIP inventory. WIP accumulates where there is a disruption to flow. The first rough mechanism to balance flow is to simply identify the points where WIP is accumulating and take measures that would open effective capacity (typically there is much hidden capacity to expose). The ongoing elaborated mechanism, which Goldratt refers to as the *process of ongoing improvement* (POOGI), involves registering the reasons where work orders do not progress as expected considering the buffer time that was consumed. An analysis

of the common reasons reveals where a focused solution will provide the biggest contribution to flow.

Turning to the environment of sales opportunities management, it is apparent we cannot apply the same POOGI mechanism. Looking at where most opportunities (WIP) accumulate does not necessarily indicate a disruption to flow, as it could be a step that simply takes much longer to carry out. Delays are certainly an indicator for disruption to flow and therefore should be an element to consider as part of the POOGI. However, in sales, unlike production, there is a much more critical indicator that should be addressed, on top of delays, for a disruption to flow—dropouts.

When designing the POOGI mechanism for the management of sales opportunities, one should take into account three different generic causes for dropouts. Dropouts could be a result of (1) a mismatch between the offer and the client—not addressing the right target market with the offer; (2) a mismatch between the offer content and the client—not adjusting correctly the offer specs to the client requirements; or (3) a faulty execution—issues in the sales process, the sales interaction with the client, the sales support deliveries, etc.

We intend to implement a POOGI on the three generic causes. A focused POOGI analysis of the third cause—faulty execution—can be done by examining the reasons for dropouts of opportunities that had a significant delay. It makes sense that an analysis of lost opportunities that were a long time in the funnel experiencing a significant delay would point to a faulty execution (if it was due to the first two reasons, we would not expect significant delays but a quick dropout). Here is how we are going to go about it:

1. We will register the reason for every delay an opportunity encounters. To determine what should be considered as a delay, we have defined the expected

standard duration of each step in the sales process. Whenever a step takes longer than the expected duration, it would be considered a delay. When this happens, the reason for this delay is registered. (We will follow the same guidelines Goldratt recommends for production—a reason would be defined as the resource or activity for which the opportunity is waiting).

2. We will focus the analysis on the opportunities that have dropped out after having a significant delay. To determine what should be considered a significant delay, we have defined a project buffer as shown in [Fig. 21-8](#). The project buffer is equal to one-third of the sales process duration. When a certain step takes longer than expected, it starts to consume the project buffer by the number of days of delay. When a certain step takes less than expected, the consumed project buffer can be recovered by the number of days gained. The project buffer is divided into three parts. If the accumulated delays consumed less than one-third of the project buffer, the status is green. If more than one-third but less than two-thirds of the project buffer is consumed (as shown in [Fig. 21-9](#)), the status is yellow. If more than two-thirds of the project buffer is consumed, the status is red. If the entire buffer is consumed, the status is black. Significant delays are considered as blacks. In other words, only opportunities that dropped out when their project buffer status was black would be subjected to this POOGI analysis.⁹

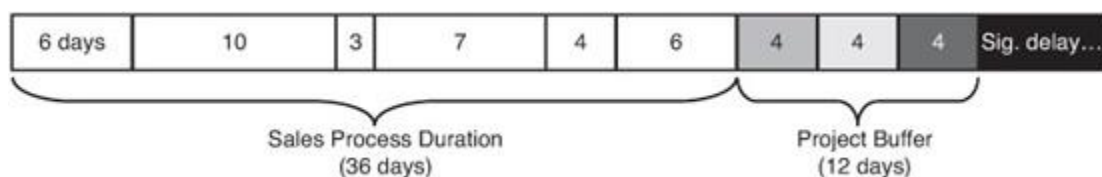


FIGURE 21-8 Sales duration and project buffer.

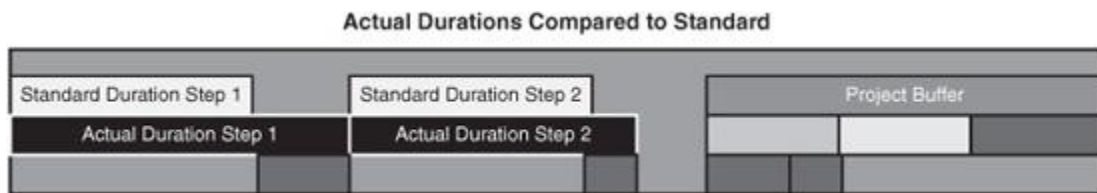


FIGURE 21-9 Disruptions to flow translated into buffer consumption.

3. We will pull out the registered reasons for the lost opportunities that had a significant delay and identify the biggest common contributor. Basically, we will identify the reason that generated the biggest accumulated consumption across all the project buffers. If the improvement efforts stemming from this analysis are effective, it will no longer be the number one contributor and another analysis will reveal the reason that should be dealt with next.

The focused POOGI for the first two generic causes would follow the same guidelines for lost opportunities that did not have a significant delay.

We expect that implementing the fourth concept of supply chain will result in another quantum jump in performance.

Third guideline: Dedicate a team to build the POOGI mechanism to identify the common significant reason for dropouts and conclude where to focus improvements efforts.

Summary

This chapter is aimed to show that what Goldratt refers to as "Supply Chain Concepts" apply much beyond what is

typically referred to as Supply Chain, and therefore should actually be referred to as the “Concepts of Flow.” Our experience in applying these concepts generated a jump in sales performance, hit ratio, and our management’s and sales team’s capabilities. In addition to the tangible results, the understanding and the application of these concepts is generating increasing harmony in the company as it becomes evident to all functions (sales, sales support, production, etc.) that they are part of one flow.

Taiichi Ohno (1988, ix) once said, “All we are doing is looking at the time line from the moment the customer gives us an order to the point we are collecting the cash and we are reducing that time line.” We humbly suggest that the underlying concepts apply much before a customer gives us an order. They apply to the same extent on our efforts to generate those orders.

Addendum

The know-how developed in the last year is substantial and probably deserves a sequel to “less is more.” Still, we thought it would be of value to give you a hint on how things look one year into the implementation of the process. Since the lessons learned do not imply any change to the solution described and only expand on it, for now we want the readers to understand how nice it is to have a different type of challenge—a challenge that can turn your sales force into becoming the real strength of the organization.

Not surprisingly, we found that applying the fourth concept of flow—POOGI—to the sales environment can lead to two paths. One path deals with disruptions to the sales flow relating to generic issues affecting the performance of all (or most) salespeople. Issues such as the sales offer design, the sales process, and the interaction with the sales support function are examples. The second path deals with

“disruptions” to the sales flow stemming from the individual performance of a salesperson.

Knowing which path one should focus on is not highly complicated, although when not applying a systematic thinking process one can easily go astray. When variability in the individual performance of members in a relevant group of salespeople is relatively low, the source of the disruption to flow probably lies in the first path. There is probably a generic flaw in one of the processes developed.¹⁰ When this is the case, trying to motivate, measure, place sanctions, or provide higher rewards to salespeople will most likely amplify frustration rather than contribute to the achievement of results. In the same manner, replacing, adding, or repositioning salespeople will really make a lasting difference only if it brings someone to change something in the flawed process and that this will work.

Identifying a systematic source for disruptions to the sales flow and removing an element that negatively affects the entire sales force can create a quantum jump in performance. One such case we dealt with last year relates to the difficulty in our environment in closing a business deal, in gaining the loyalty of a client, and with it the majority of its business. How can one win the business (almost every order) of a client when every order is a new product that requires development, when the client is compelled to obtain quotes from different suppliers for every order? Overcoming this challenge, this systematic source of disruptions, required a change (or an addition) to our marketing offer.

The second path, the one relating to the individual performance of a salesperson, as trivial as it may seem, makes us aware of the fact that “we are dealing with human beings.” Sales people have different skills, motivations, ambitions, and learning curves. Not all care for the same things. Some parts of the execution process may be more

natural to a particular person; some clients may be more suitable for a specific type of personality. If we want to dominate the complexity of the elements of the process and the interactions involved in closing a business deal, the management and guiding of the individuals in the sales force becomes a key element.

The paradigm shift comes when we clarify the conflict between dealing with the sales force in the traditional way, to “show them their low performance and to put pressure on them,” and dealing with them “on specific parts of the sales process.” The tendency is to think that because they are salespeople, they must know what they are doing is completely wrong. Just think about how most sales managers are conditioned to deal with their sales force. Isn't it true that the salespeople know their results? Personally, we believe that good salespeople want to sell, whether or not they have commissions. When they do have commissions, no one would argue that not selling brings a lot of pressure without the managers doing it.

Assuming the main generic disruptions to the sales flow are removed (the first type), a significant jump in performance can be achieved by treating the sales team as “professional sports athletes.” Make them improve their bad shots and use more of the good ones.

Managing the sales force with a “system” allows us to identify the steps individual salespeople fail to execute well, and to devote time together with them to understand why it happens. In this aspect, the logical thinking tools of TOC have played a major role. For example, many challenges a salesperson has, and probably every deviation from a process a salesperson makes, can be analyzed, together with him or her, by using the conflict analysis tool presented earlier in this chapter (the “Cloud”).

In the past year, we have learned that applying the POOGI to the sales flow enables one to focus on the generic disruptions, the ones that affect the performance of the

entire sales force, while dealing with individual performance gaps. It is the “system” view—the one focusing on flow—that enables such a powerful lever.

References

Goldratt, E. M. 2008. The Goldratt Webcast Program on Project Management. Roelofarendsveen, The Netherlands: Goldratt Marketing Group.

Goldratt, E. M. 2008. “Standing on the shoulders of giants,” White Paper presented at TOCICO International Conference, Las Vegas, NV.

Goldratt, E. M. 2009. “Standing on the shoulders of giants,” The Manufacturer. June. http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants. (accessed February 4, 2010).

Goldratt, E. M. 2008. Strategy & Tactics Tree for Make-to-Order Companies. These can be found in the S&T Library embedded in HARMONY (S&T Expert System) downloadable from www.goldrattresearchlabs.com.

Ohno, T. 1988. *Toyota Production System*. New York: Productivity Press.

About the Authors

Mauricio Herman is presently CEO of a company that offers integrated product solutions from designs done to customer requirements, to packaging and display in retail stores. He has worked in various departments across his

company, learning operations from the bottom up. Mr. Herman has completed the “Jonah” course in Theory of Constraints and has followed TOC concepts in areas of his business including Sales and Finance. He has coauthored a leading article on the application of TOC to the Sales Process, which formed the basis for this chapter.

Rami Goldratt is the CEO of Goldratt Consulting—the front edge TOC consulting firm, funded and chaired by Dr. Eli Goldratt.

Rami is recognized worldwide as a leading figure of the TOC Body of Knowledge. After serving as Goldratt Consulting head of development, Rami was appointed CEO. He was the CEO of SFSCo (Solutions for Sales Co.), the specialized supplier of Sales and Marketing experts for Viable Vision projects—a holistic solution based on TOC implemented with companies worldwide. He received his MA in philosophy from Tel Aviv University, Israel.

CHAPTER 22

Mafia Offers: Dealing With a Market Constraint

Dr. Lisa Lang

Spend just two hours reading this chapter and if you don't get at least one good idea for your business, contact me and I will give you a refund!¹

—Dr. Lisa

That's a Mafia Offer² and it's real. The purpose of this chapter is to introduce you to the Mafia Offer—the *Theory of Constraints* (TOC) marketing solution. The chapter progresses from the discovery of what a Mafia Offer is, to the guidelines for creating an offer, to how to present an offer, and ends with how the reader can create their own Mafia Offer.

The best way to read this chapter is in order. Each section builds on the subsequent sections, so that when you reach the summary you will have a very good understanding of what we in TOC call the “Mafia Offer.”

Introduction: What Is a Mafia Offer?

A *Mafia Offer* sounds like something out of a movie, not something that could seriously help you make more money in your business by increasing and controlling your sales.

Dr. Goldratt first introduced the concept of a Mafia Offer in his book *It's Not Luck* (Goldratt, 1994, 133). Later, he defined a Mafia Offer as “an offer they can’t refuse” (Goldratt, 2008, 67). But in writing, he more frequently refers to it as an *unrefusable offer* (URO; Goldratt, 1999, 120) and more recently he (and Goldratt Consulting) emphasizes the need to establish, capitalize, and sustain a decisive competitive edge (Herman and Goldratt, 2008).³

Copyright © 2010 by Dr. Lisa Lang.

For this chapter, I will use the term *Mafia Offer* and have defined it as follows: *An offer so good that your customers can't refuse it and your competition can't or won't offer the same.* In addition, I will refer to the operational improvements required for a Mafia Offer as the decisive *competitive edge*, operational advantage, or competitive advantage.

A Mafia Offer is simply the offer you make to your market—your prospects and customers—to make them desire your products or services and something that your competition cannot quickly match. And, of course, the offer you make is a combination of your products, services, and how you deliver them. Moreover, for your offer—the solution you’re selling—to be *unrefusable*, you are most likely offering something of equal or greater value than the price you are charging.

Many people confuse a Mafia Offer with a *unique selling proposition* (USP), *customer value proposition* (CVP), or a *sustainable competitive advantage* (SCA).

At first blush, it would seem that a Mafia Offer is similar to these other terms; however, when most people are talking about these alternatives they are actually quite different from what TOC experts mean by a Mafia Offer.

USPs, CVPs, and SCAs take what you already do and state it succinctly and with more specificity aimed at one or a few

of their customers' problems or gaps in current market offerings. These alternatives can be Mafia Offers, but most of the time they are not. Furthermore, an SCA is, in my view, an operational or technological (although these are *not* typically sustainable) advantage and not an offer per se.

Most companies offer solutions that solve their customers' various problems or symptoms. With a Mafia Offer, we are addressing our customers' core problem as it relates to doing business with our industry.

A Mafia Offer typically requires that you do something different (make operational improvements to establish a decisive competitive edge) to address your prospect's core problem. These operational improvements allow you to actually deliver something unrefusable to your customers and something that your competition can't or won't do because they are not willing to or don't know how to make the same improvements. In other words, you have to establish an operational advantage.

In this way, a Mafia Offer is a sustainable market offer built on this advantage. Mafia Offers are not a positioning or a tag line and "can only be created by satisfying a significant need of the market to the extent that no other significant competitor can" (Herman and Goldratt, 2008).⁴ A Mafia Offer is where we start if you have a market constraint.

Do You Have a Market Constraint?

Let's do a quick check. How would you answer this question? If I could increase your sales tomorrow by 20 percent, could you handle the increase while:

- being 100 percent on time, to your very first commitment;

- without going into firefighting mode;⁵ and
- still maintain a competitive lead-time?

If the only way that you could handle the increase is to increase your lead times, work overtime, or miss due dates, then you have an internal operational constraint. You don't have a market constraint. On the other hand, if you can answer yes and you could take a 20 percent increase in sales and not have any negative effects, then you do have a market or sales process constraint.

To determine if the issue is a market constraint or a sales process constraint, we need to determine how you would answer another question—why should I buy from you?

Imagine that you just walked into a hot prospect's office and the prospect said, you're the third vendor I'm interviewing today so let's cut to the chase and just tell me—why should I buy from you? Why should I choose you over the others?"

Before you continue reading, please write down the reasons *your* customers and prospects should buy from you. Make a list.

I've asked this question around the world and have collected many answers along the way (Smith, 2006, 100).⁶ However, there are themes that tend to repeat. Most people answer the question with something like this:

- We have outstanding quality and it's better than the competition.
- We have a great reputation.
- We get good results for our customers.
- We have very knowledgeable, great employees with low turnover.
- We're very responsive.
- We're very innovative, helping our customers to ...

- You can trust us.

Some version of that is what most people write on their list. The list might vary slightly depending on your industry. However, it typically does not vary much between you and your competitors. And that's really the point.

If you're saying the same things as your competition, then you're not really providing any compelling reasons. Therefore, you sound the same as your competitors.

Moreover, what do you typically do if you have heard something before? Do you sit up and listen closely, hanging on every word? Well, if you are like me, you tune out. They might as well be saying blah, blah, blah ... because that is what you are hearing.

Therefore, if I'm a buyer and you and your competitors are saying some version of the same thing, then I might as well choose who to purchase from based on price. In addition, even if you are truly different or better than your competition in some way, it doesn't really matter if your prospect doesn't get it. A cute little tag line⁷ is not likely to change my mind or help me to get it.

As long as you sound the same as your competitors, we must assume that your constraint is a market constraint. In other words, if you have not convinced your prospects that what you provide is of greater value relative to the price you are charging and of greater value than the competition, then why would they buy from you?

Only after you have a good offer, a Mafia Offer, and you are delivering it correctly and your sales are still not increasing, can we determine that you might have a sales process constraint.⁸ Therefore, we need to start by creating a Mafia Offer. A good Mafia Offer delivered correctly is the solution to a market constraint.

Developing a Mafia Offer

To develop a Mafia Offer, there are three things that we need to consider.

1. Your Capabilities, Both What They Are and What They Could Be, Compared to Your Competition.

Your capabilities are how you deliver your product or service. For example, what's your lead time? Your due date performance? Your quality? Your answers to these questions are your capabilities. Typically, when we first start working with a company, their capabilities are similar to those of their competitors. If they were much better or much worse, they would know.

If you're quoting a 6-week lead time, typically your competition is also quoting a 6-week lead time. Your prospects make sure that you know if someone is quoting a better deal. So typically, everyone in a niche quotes similar capabilities to ensure they don't lose opportunities. Moreover, of course they would quote better if they could actually deliver better.

To determine what your capabilities could be, we rely on experience with the TOC logistical solutions. To give you an idea of possible results, see [Fig. 22-1](#). From [Fig. 22-1](#), if you were quoting a 6-week lead time, we can expect a 70 percent reduction or a lead time of less than 2 weeks after applying the TOC solutions to your operations.⁹ This gives us an idea of the decisive competitive edge that we can establish and capitalize on in our offer.¹⁰

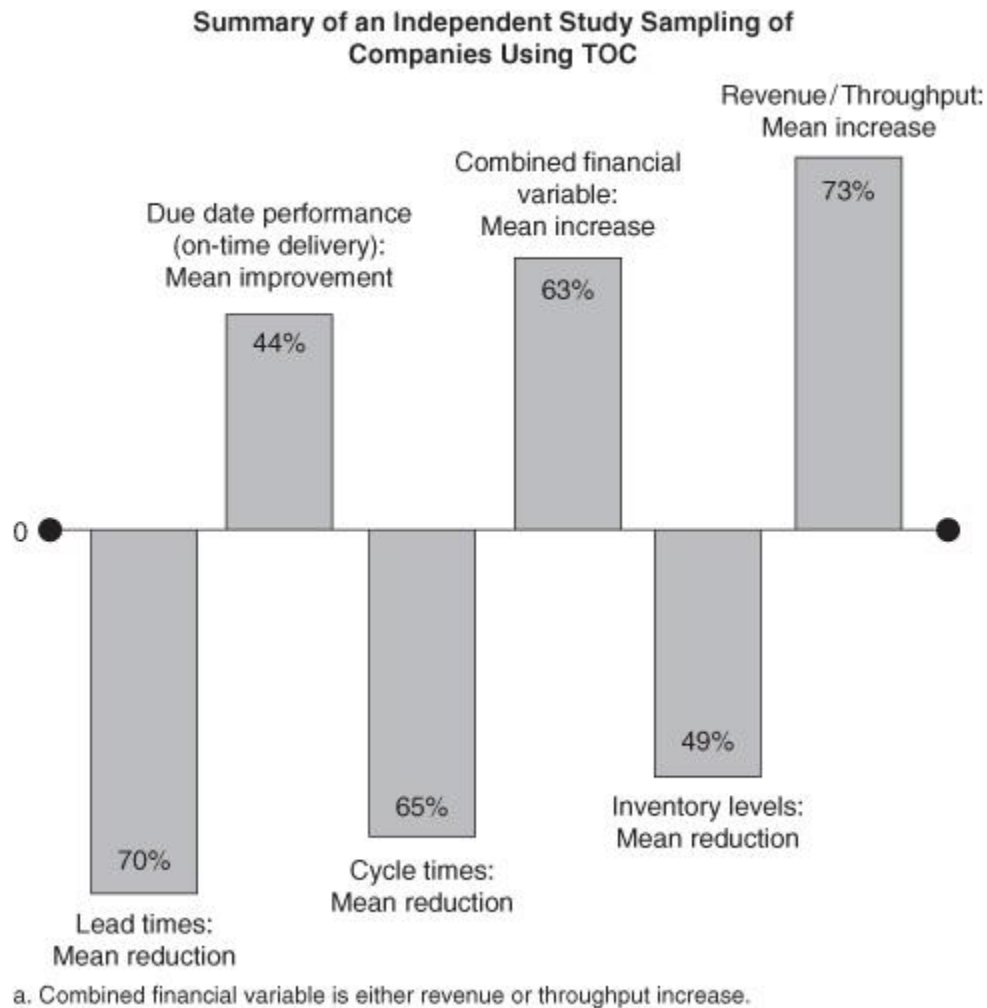


FIGURE 22-1 Typical results with Theory of Constraints. (Source: Mabin and Balderstone, 2000.)

2. Your Industry—How You and Your Competitors Sell Whatever You Sell.

The second thing we look at to develop your Mafia Offer is how your industry sells whatever it is that you sell. A whole slew of questions will fit for your industry. Here are some examples that may or may not apply to you:

Is it industry practice to use a price/quantity curve? How do you and your competitors' typically charge? By

the hour? By the day? By the project? Time and materials? Flat rate? Who pays for shipping? Paid at the start? Paid at the end? Progress payments?

The key is to understand how your industry interacts, in the selling and delivering of your products/services, with your typical prospects and customers.

3. Your Specific Customers and How They are Impacted by Typical Capabilities and How Your Industry Sells.

Since your customers are the only judge of your Mafia Offer, we also need to understand how your current capabilities and those of your competitors affect the companies in your target market, and how they are affected by the way you all sell to them.

It is in these interactions and interfaces that we may be causing negative effects for our customers and prospects. Understanding these negative effects leads us to uncover our customer's core problem relative to doing business with our industry.

The easiest way to understand what a Mafia Offer is, what makes it good, and how to create one is to go through an example. Most likely, this example won't apply to you because an offer is specific to a company and its particular customers. Nevertheless, you can gain by understanding how to apply the three considerations to a specific situation.

Custom Label Printer—An Example

Let's consider a custom label printer. The labels this printer makes for one customer can't be sold to anyone else. However, the same customer may reorder a label for a number of years. Moreover, this label company's customers buy 100+ different labels for their various products. Many of their customers are regional-sized food and beverage

companies. They produce food products in multiple flavors and put them in a variety of packaging, causing them to need 100+ different labels.

The analysis started by evaluating the internal capabilities of this printer and those of their competitors. We found that this printer and their competitors generally quoted a 2-week lead time. We also learned that the due date performance was about 90 percent for this type of custom label printer.

Determining operational performance of the label company was straightforward. The only thing you need to watch out for is how they calculate *due-date performance* (DDP). Some companies will change the due date commitment if they call the customer and get permission to be late. If they get this permission and meet the new date, they consider this to be on time. So we ask them to calculate their DDP on “first date given.”

To determine if this label company’s performance was typical for their competitors as well, we simply talked to the salespeople. If a company’s performance is much better or much worse than the competition, their salespeople will have heard about it. If you don’t have salespeople, then whoever the chief salesperson is (like the company owner) is the one to ask. In this case, the salespeople indicated that the 2-week lead time and the 90 percent DDP was neither praised nor a problem.

From our experience of working with printers in the past, we expected to get the time through the shop down to just a few days and to improve DDP to 99+ percent. A quick tour through the shop verified that this should be possible. On the tour, we noticed a large amount of *work-in-progress* (WIP). We also noticed one of the printing press operators going through the pile of jobs in house and asked what he was looking for. He said that he was trying to determine the best way to “gang the jobs” to make the best use of his current setup.

With this information and knowing that the actual touch time of a job was measured in minutes, we were confident that the cycle time through the shop should go from the 2 weeks or more now to just a few days. We anticipated that we should be able to improve this shop to 99+ percent DDP and the time it takes an order to flow through the shop should be dramatically reduced. Our estimate was two to three days.¹¹

Next, we turned our attention to the industry and how the industry sells custom labels. Now, if you've ever bought anything printed you know that the lower the price per piece that you want, then the more you will need to buy. If you only want one or a few pieces, then your price per piece will be very high. The printing industry uses a price per quantity curve like the one shown in [Fig. 22-2](#).

In addition to a price per quantity curve, it was also standard practice for this label printer and its competitors to allow customers to spread the quantity across all their different labels. Therefore, if a customer needed 100 different labels they could spread the volume across all 100.

Next, we looked at the impact that the 2-week lead time with 90 percent DDP along with the industry practices have on the label company's customers. In other words, what negative effects for our customers are we causing because of our capabilities and how we sell?

In the case of this custom label printer, we selected a representative customer to understand the cause-and-effect relationship between how we sell and the impact it has on our customers. We selected a coffee roaster that purchased about 100 different labels and who, when they looked at that price per quantity curve, decided to purchase 6 months worth of labels at a time. Labels are relatively small and inexpensive, so holding 6 months of inventory was common.

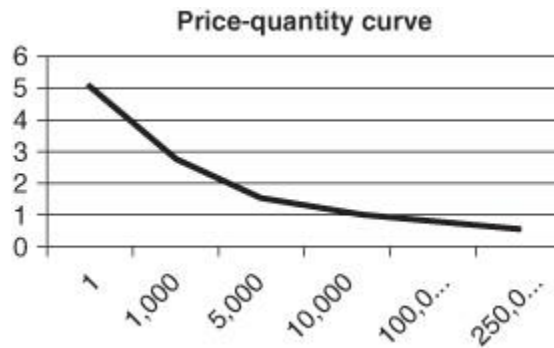


FIGURE 22-2 Price-quantity curve

To finalize the order of six-months worth of labels, the coffee roaster needs to determine how to spread the quantity across the 100 labels. How many French Roast, Columbian Roast, and French Vanillas were going to sell in each size bag? To do that they had to forecast out 6 months how many of each label they were going to need, which meant they had to guess:

- how much coffee all of us were going to buy;
- in what flavor; and
- in which size bag.

Now, if you only know one thing about a forecast, what do you know? That it's wrong! The only question is just how much is it wrong and in which direction?

Our practices cause our customers to have to forecast. What are the negatives that our customers would be experiencing from a wrong forecast? This is easy to check. I went into the customer service department of the label company and asked the customer service people two questions:

1. Do you ever get frantic calls from customers who have stocked out of labels? They said, "Yes, we get

those calls all the time.” What’s “all the time”? They indicated they were getting 2 to 3 of those calls a week!

2. Does the opposite also happen? Do you have customers who typically order 6 months worth of labels, but it’s been over 6 months since some of the labels have been reordered? Customer Service responded, “Yes, that also happens. In fact, the coffee roaster you were just asking us about called last week. They were frantic because they were out of Columbian Roast labels for their one-pound bags. And while we had them on the phone we asked them if they also wanted to order French Roast labels because it had been over 9 months.” They responded, “We have enough French Roast labels for our grandchildren, so just send the Columbian Roast!” The customer explained that their lines had gone down when they ran out of labels and they needed them ASAP to fulfill an order. They asked the label company to ship them overnight.

Our practices force our customers to forecast. The forecast ends up being wrong in one direction or the other. If the forecast is low, their lines go down, causing them to lose productivity and to work overtime when they do finally get the labels in. Their costs also increase because, in addition to the overtime, they have to pay expedited shipping charges. In addition, the buyers are frantically working to get the labels in house and the line back up.

If the forecast is high, they end up with too much inventory of some labels. High inventory levels increase the likelihood of damage or obsolescence. In addition, high inventory results in higher carrying costs and cash tied up in unneeded inventory, and causes the company to hesitate before making any label changes.

Therefore, our analyses¹² lead to the following Mafia Offer:

“Mr. Customer, don’t give me orders. Your orders are based on your best guess of how many labels you think you might need. That’s because label printers put that price per quantity curve in front of you and force you to have to guess out six months. The forecast ends up being wrong, and how can it possibly be right? Instead, tell us every day how many labels you use and we can guarantee, on the one hand, that you won’t have to hold more than two weeks’ worth of labels. And you know how your marketing department was complaining that they can’t make the changes they want because you have six months worth of inventory? Well, now you will only have two weeks. At the same time, we will guarantee that we never stock you out. We will guarantee that you’ll never go to the shelf and not have the label you need. And if we ever do stock you out, we will pay you \$500 per day per label. We offer all this at the same competitive price you pay today and of course you will have a lot less of your cash tied up.”

The Test—Is It a Mafia Offer?

Let’s test that offer against our definition. Is the offer so good our customers can’t refuse it? Well, that depends on the customer. If we have done a good job with our analysis, it should be unrefusable to 80+ percent of the target market. Realize that no offer will be 100 percent accepted by any market. There will just be some people, for whatever reason, that won’t find your offer compelling.

“Reason is not automatic. Those who deny it cannot be conquered by it. Do not count on them. Leave them alone.”

—Ayn Rand

When we develop a Mafia Offer, we start by asking to whom will the offer be made? We select a target market—a type of customer. The market we select can depend on a number of issues; for example:

- What market do we want to grow?
- What market has the best margins?
- Do we have too much business with one customer or in one market?
- Which customers or types of customers do we dread? (If our competitors also dread these customers, they may be more easily acquired.)
- What market has tons of room for us to grow?

However, the key is that our analysis is done with this target market in mind. In our example, most of the label company's customers were regional-sized food and beverage manufacturers. The offer was developed for those customers and prospects.

Equipment manufacturers also purchase labels. However, this offer would not work for them. They typically know that they are going to produce 100 machines this year and they know they will need 500 labels for those 100 machines. They do not have a forecasting problem to the same extent that food and beverage manufacturers do. They would not likely be moved by our offer, so our prospecting attention would be better spent on food and beverage manufacturers who struggle to keep the correct mix of label inventory while still having a mountain of inventory.

So why is the label company offer unrefusable to food and beverage manufacturers? Let's make a list:

- It reduces their inventory from about 6 months to 2 weeks.

- It reduces the amount of cash they have tied up in inventory.
- It eliminates the chaos that results when a stockout occurs.
- It reduces the costs associated with stockouts—down time, expedited shipping, and overtime.
- It reduces the inventory carrying costs.
- It reduces inventory obsolescence and there are fewer labels to damage if an incident should occur.
- It provides increased marketing opportunities when you can quickly make changes.
- It eliminates the need to place orders and do forecasting, therefore freeing up that time for other activities.
- And, all that is realized for the same price.

Therefore, we can conclude that this offer is unrefusable to our target market, but can our competition match it? We are asking our customers to hold 2 weeks' worth of inventory, down from about 6 months. What's the competition's lead time? If you recall from our analysis, the standard lead time was 2 weeks with 90 percent DDP. Therefore, there is no way our competitors could match the offer and not have to pay penalties or to hold a substantial amount of inventory at their risk.

As it turned out, we improved our flow from over 2 weeks to just 2 days (while sales and staffing remained constant), establishing the basis for a nice decisive competitive edge. Therefore, we should never have to pay a penalty as long as we are paying attention and we know how to react to the daily consumption data. So, this offer does meet the two requirements for a Mafia Offer. It is an offer the customer can't refuse and the competition can't offer the same.

What Did It Take to Make the Offer?

In addition to improving operations by implementing *Simplified Drum-Buffer-Rope* (S-DBR),¹³ the label company had to change their thinking in a number of areas. First, their offer would require that they do more setups. Ask any label printer how much it costs to do a setup and they will tell you to the penny. However, how much does it really cost?

Nothing. You don't pay your employees by the setup, and you don't pay your machine by the setup. The only real cost is a little paper and ink to get everything lined up. This is so small and so hard to allocate an exact cost perfectly, I just think of it as nothing. However, the label company's competition thinks that there is a real cost and even if they could match the offer, *they don't want to!* They think the label company's costs will increase and they will go out of business.

The whole reason the industry uses a price per quantity curve is to save these setups. However, saving setups is about printing, about our costs, it's not about the customer. In fact, our analysis showed that the price per quantity curve leads to the need for our customers to forecast. And that leads to a number of negative effects.

So one of the biggest changes the label company had to make was in how they think about their costs. They had to understand that the true cost to do more setups was practically nothing and saving time on a non-constraint would save nothing. Setups do take more time, but an interesting thing happens when you start to do something more often—you get better at it! The label company freed up capacity by not wasting production time making labels that were not needed. So despite the additional setups, flow through the label company stayed at about 2 days.

So there you have it, an offer that is so good our customers can't refuse it and something the competition can't *and* won't match! The competition will not match this

offer for some period of time and maybe never. Therefore, we've built and capitalized on a very sustainable competitive advantage.

A Mafia Offer Is NOT . . .

Mafia Offers do not require an innovation. There is the innovation camp that believes the only way to gain substantially more sales is to innovate better and faster, the stuff your customers want. Some have even gone so far as to call your existing products, in your existing markets, a bloody red ocean due to all the fierce competition.

In the book *Blue Ocean Strategy* (Kim and Mauborgne, 2005), the authors contend that it is *not* possible, in most cases, to sell more of your existing products in your existing markets. They lay out a process for developing new products for new markets—a highly risky endeavor.

Innovation is absolutely necessary for long-term sustainability, no question. My issue with using innovation as your sole means for increasing sales is that it's short lived. How long do most innovations last? How long does it take your competitors to copy?

If new products and new markets are a risky proposition, why use innovation as your sole approach to increasing sales? And why would you take such a risky action if you could develop a Mafia Offer? The answer, of course, is you wouldn't. Nevertheless, innovation is the only logical alternative if you are not aware of TOC or Mafia Offers.

I would say the same thing about price. Price reductions can also be copied very quickly and do not typically provide a sustainable advantage. Therefore, Mafia Offers are not solely based on price.

In addition, remember this list:

- We have outstanding quality and it's better than the competition.
- We have a great reputation.
- We get good results for our customers.
- We have very knowledgeable, great employees with low turnover.
- We're very responsive.
- We're very innovative, helping our customers to ...
- You can trust us.

These are *not* the qualities of a Mafia Offer. Your competitors say the exact same thing. All good companies have these qualities or they wouldn't be in business for long. A Mafia Offer is *not* a list of strengths, a cliché, subjective, or offered by the competition.

In the custom label company example:

- The Mafia Offer was developed for a regular company with its *existing products* in its *existing markets*.
- This company had no particular competitive advantage or innovation—no patent, no unique technology, the same equipment as competitors, and similar employees.

Yet, it was not based on a price reduction, not easy for the competition to follow, and so good that most customers would accept it readily.

Where to Start?

Should we improve operations or create a Mafia Offer first? We've done it both ways and either way works. But I like to start by developing the Mafia Offer. Once you have the offer, you know to what degree you need to improve operations.

More importantly, it gives you a reason to change. We have found that when we start with the offer, the client is more motivated to make the operational improvements. The operational improvements occur faster.

When we have worked with clients that have already implemented TOC in their operations, we often find that they have been giving away some or all of the improvements they've made. This is particularly true in cases where they had very bad DDP. They improve operations and then they give away the shorter lead times because they feel guilty about their past performance. Therefore, my recommendation is to create your Mafia Offer first, and then make the operational improvements necessary to deliver your offer.

However, before you make your offer there are some things you must do. Getting your operations in shape is paramount. The fastest way to kill a Mafia Offer is to not be able to deliver it. So, make sure you can deliver your offer by doing a couple of dry runs. Pretend that several of your orders or jobs are for your offer and see how you do. Alternatively, if you are going to guarantee a shipping date, determine how much you would be paying in penalties if every order were guaranteed.

Making sure your operation is ready for the offer is straightforward. However, you should also predict what could go wrong from both your perspective and your customers' perspective. This will help you to determine if you've missed anything and to create some of the details of your offer.

This is *not* permission to create a bunch of small print (weasel words) for your offer. The objective is to predict negative branches and to trim them. When in doubt, do not add weasel words; instead favor your customers' position.

Protecting yourself and looking out for your interests is what caused the negatives for your customer in the first

place, so don't back track. But at the same time, don't put your entire business at risk.

Sustaining the Advantage and the Offer

Based on experience, a good Mafia Offer will give you years on the competition. The competition thinks your offer is going to put you out of business. It will take them some time before they even take a second look at what you're doing.

In this chapter, I have laid out a very nice Mafia Offer for a custom label printer. I have been using this example for years. I have been the keynote speaker at the Tag and Label Manufacturers Institute annual conference twice. And despite that, no direct competitor has copied this offer. It is true that the offer may not work perfectly for another label printer because it has different customers. However, at least some of it would be transferable. So why don't they do it? Why don't other industries who have similar industry practices with similar negative effects on their customers give it a go?

First, I think the "we're different" thought keeps us from going too far with it. Then, even if someone starts to look into it, they can get blocked in any number of ways. Developing and implementing the necessary changes for a Mafia Offer requires multiple paradigm shifts. In particular, it requires that you change the way you think. And changing the way you think about costs, setups, multitasking, WIP, scheduling, and how one is supposed to go about making money is very difficult. To make all those changes at once is even harder.

Most Mafia Offers, by definition, are not easily copied by your competitors, but another source of sustainability problems is you. Once you start making your offer and your sales start to increase, there are some potential negatives to success. The easiest way to avoid these negatives is by

using TOC techniques and measures. Here are some watch outs:

- Your load relative to your capacity has increased, due to the increase in sales, which causes you to:
 - start missing your commitments
 - increase your quoted lead-times
 - speed up operations or cut corners causing your quality to decrease
- The interest in your products/services is much higher than before the offer, so
- Some leads are starting to slip through the cracks
- Your customer service starts to decline
- The first step of your process, which is required for new customers (like design or engineering), has become a bottleneck

If you continue to measure and pay attention, you should be able to avoid these problems. They are certainly predictable and if you pay attention to your load versus your capacity (in all parts of your operations and sales processes), then you can make the necessary preparations and responses.

Benefits to the Label Company

We previously discussed why the Mafia Offer was unrefusable for the label company customers. But what are the benefits for the label company?

- They stop the blah, blah, blah and sounding like their competitors. They can answer, “Why should I buy from you?”
- Sales increase (and so do profits if the TOC logistics solutions are used to improve operations).

- Time wasted producing labels that are not needed is eliminated.
- They gain 100 percent supply for label stock-keeping units (SKUs) included in the program.
- They substantially reduce the risk of losing a customer to a competitor over a small price reduction. Customers taking advantage of the Mafia Offer ask for long-term contracts.
- They became better at doing setups, and can easily run small batches, increasing their flexibility and responsiveness to the market.
- They became very good at adding new customers.
- Cash flow improves due to smaller batches and more frequent billing based on replenishment that is more frequent.

It's a Business Deal

Many Mafia Offers are business deals, and as such, are sold differently. The label company is no longer selling labels; they are selling the guaranteed availability of labels based on their customers' actual use or consumption. For that offer to work, the label company will need to supply 100 percent of the labels that are included in the program. In addition, the customer will need to supply daily consumption data for those labels. This transfer of data sounds much scarier than it actually is. Typically, an electronic transfer can be set up to occur automatically each day.

However, the point is that the supplier (the label company) and the customer are more closely integrating. Both sides stand to benefit from this business collaboration.

The offer needs to be presented in a way that gets the customer engaged, interactive, and ready to buy. The way to do this is very different from what salespeople do today in a typical sales call.

The biggest issue we see *after* someone already has a good Mafia Offer is in how it is delivered. So let's talk about that. How should a Mafia Offer be presented? We need to get this right because a good offer, delivered poorly, won't increase sales.

We have already discussed what happens when you go into blah, blah, blah mode—your prospects stop paying attention. Therefore, we need to present our offer in a way that is compelling, gains their trust, and gets them to take action. To improve on my ability to successfully present Mafia Offers and to help my clients successfully present their offers, I've studied and applied some basic psychology. This psychology, combined with the TOC buy-in process, has led to the success we have had with Mafia Offers and marketing in general.

The Psychology of Delivering a Mafia Offer¹⁴

Neuroscience, using a technology called functional MRI, has helped us to understand what part of our brains is involved in making decisions. The outer-most part of our brain (the newest or youngest part) is where rational thinking takes place. The middle part of our brain gives us our gut feelings and all the emotional components related to making a decision. Nevertheless, the decision maker is the core of our brain. This core is the oldest part of our brain and has been called the *old brain*, the *reptilian brain*, the first brain, or the limbic system. It doesn't matter what you call it, what matters is that we use (and our *prospects* use) the most ancient part of our brain to make all of our decisions.

Brain scientist Leslie Hart determined that the old brain is the part of our brain that decides what senses get transferred to the new brain, and more importantly, what decisions will be accepted. (Hart, 1975).

This means that we must better understand how the old reptilian brain makes decisions for us to sell successfully. There is good news and bad news with this. The bad news is that our prospects and we are making decisions at the primitive level of a crocodile or frog. The good news with the old reptilian brain is that it's so ancient and so primitive that it becomes predictable; it's been estimated that the old brain is approximately 450 million years old (Ornstein, 1992). Therefore, if we can understand how to predict what the reptilian brain will do, we can better sell to it.

According to Renvoise' and Morin (2007, 11), the old reptilian brain, besides processing input from other parts of our brain, only responds to six stimuli. Those stimuli are:

1. Self-centered—It's all about me and my preservation.
2. Contrast—Say the same thing I've already heard and I tune out. Say or do something in contrast and you have my attention.
3. Tangible input—Simple, straightforward is best.
4. The beginning and the end—To conserve energy, the old brain may stop paying attention in the middle.
5. Visual stimuli—Visual works best with the old brain.
6. Emotion—Emotion rules. We are not thinking machines that feel, we are feeling machines that think (Damasio, 1995).

Therefore, if we can understand how to apply the use of these six stimuli, we have the key to engaging our customers/prospects in our Mafia Offer. In addition, if we

combine this with the TOC solution for sales (Goldratt and Goldratt, 2003)¹⁵ and buy-in processes, they may actually decide to buy from us.

So let's review the buy-in process in light of these stimuli. The buy-in process has evolved over time and you can find different versions of it. I'm going to review the steps we typically cover when presenting a Mafia Offer and how we might do them keeping the six stimuli in mind.

Agree on the Problem

Since the old reptilian brain is self-centered and concerned with its own survival above all else, it is highly interested in solutions that will alleviate any pain it's feeling or problems with which it's dealing. That is why humans spend more time and energy avoiding pain or looking to destroy pain than we devote to gaining higher levels of comfort.¹⁶ Focus on the problems and pain your prospect is experiencing, not the features of your products or service.

Which magazine do you think men are more likely to buy?¹⁷ A men's health magazine with the cover, "Lose Your Gut Fast" or a similar magazine with the cover, "Get Six-Pack Abs"?

One study showed that over 80 percent of men chose the first cover—"Lose Your Gut Fast." Why? People are more interested in avoiding (or reducing) pain than they are in increasing pleasure.

Agree on the Direction of the Solution

Have you noticed that a large portion of all Websites and brochures start with the same sentence, "We are one of the leading providers of ..."? Or, they have a picture of their

building on the home page. If you're sitting in a presentation, have you noticed that most start by the presenter expounding on the history of the company?

This blah, blah, blah is the typical way most people approach their market. Such empty claims, neutral statements, or general filling of silence work against you. To reach the old brain you should say (and prove) a contrasted statement because the old brain responds favorably to clear, solid contrast. Powerful, unique solutions attract prospects because they highlight the difference, gap, or disruption the old brain is proactively looking for to justify a quick decision.

Agree the Solution Solves the Problem

Focusing on the unique benefits of your solution is all well and good, but technically it doesn't prove anything. Remember, the old brain prefers tangible, simple, straightforward information over complicated or abstract concepts. It needs solid proof of how your solutions will enable it to survive or benefit. Since the old brain can't decide unless it feels secure, you need to concretely demonstrate, not just describe, the gain your prospects will experience from your product or service—the results of a specific solution to their problem—in a way that satisfies the old brain's need for concrete evidence. So it's not just about value, it's about proven value or proven risk reduction.

This has implications not only for our Mafia Offers, but also for how we approach prospects in our emails, Websites, and brochures. There are implications on how we describe the problem, how we describe our Mafia Offer, and how we agree or prove that our offer will provide the results.

So let's walk through the main components of the label company's Mafia Offer solution for sales presentation and

how it might be delivered in light of what we now know about the old reptilian brain.

Here's the typical flow.

We *never* start with who we are and how long we've been in business and all that typical blah, blah, blah stuff. The old brain doesn't care. We start with something like:

“We did an analysis of our industry. We looked at our practices and the practices of our competitors. And we discovered that our practices are having a negative effect on *your* bottom line. We would like to share and check that analysis with you.”

In this way, our opening statement (*the beginning*) is about them (*self-centered*) and *their* bottom line.

Also notice that this is in *contrast* to what most of their suppliers do.

Agree on the Problem

In the PowerPoint presentation, we start with “Analysis of the Suppliers’ Practices.” In this part of the presentation, we show how suppliers in our industry (our competitors and us) have a negative impact on our customers’ business. These negative effects are due to our practices. Typically, these practices are common across our industry and include minimum order requirements, scheduling practices, lead times, and so on. In this way, we are starting with how our practices are the cause for at least some of their problems. Typically, we will do three slides like the one shown in [Fig. 22-3](#), showing the negative effect our practices have on our customers. Then we summarize in one slide. To deliver to the old brain, we stress that these industry practices are having a negative effect on *their* business. Again, make it about them.

You also can make the problem visual by adding a picture of mountains of inventory. I also like to generate discussion around these problems because oftentimes there are people in the room who were not aware of the situation or the magnitude of the problem. I want to get them a little emotional about the pain. I might ask, for example, if in fact they have had the experience of having to hold higher inventories due to a supplier's policy.

I can often get someone to tell a story, and if he or she does, I try to make it tangible by asking how much, how big, or whatever the appropriate question might be. When you do the typical sales call—the “show up and throw up” approach—spouting all the features and benefits of your product or service, the customer is automatically resisting and looking for reasons not to buy. By starting with how we negatively affect them, customers are more open to hear what we have to say next.

Supplier Practice: Minimum order quantities and volume discounts

Customer's mode of behavior: Batch orders (delay ordering) to accumulate needs and order larger quantities to get the discount.

Implications on your business: High inventory with all the cost and risk that goes with it. And lots of cash tied up!

FIGURE 22-3 Analysis of suppliers' practices.

I can't stress enough how important it is to really nail this first part of the presentation. In no more than four slides and 10 minutes you should be able to describe how *your* industry (you and your competitors) are having a negative impact on your customers' bottom line. If you can do this instead of the typical “Background of Our Company” and

“Background of Our Products,” they will be eager to hear what you have to say next, instead of being half asleep.

However, if your analysis was *not* correct and you have not correctly identified the pain, then you are *not* going to make a sale, so thank them for reviewing your analysis and leave.

However, if we have our prospect’s head nodding in agreement and they have shared a couple of stories, they are actually eager to hear what we have to say next. You are the first vendor that has so eloquently described the dynamics between industry practices and you verbalized it better than even they have or could.

Agree on the Direction of the Solution

Usually I transition by saying something like, “So, if we have accurately captured the problems that our industry causes you, we then need to determine criteria for a good solution.”

We then present and review the slide with this criteria and get the prospect’s feedback. We also note that this criterion should be used to evaluate any potential solution, even one from a competitor. This is so we can create *contrast*.

Next, we ask what it would be like if you had a solution that met that criterion. Here we are creating a *vision* and tapping *emotions*.

Once we have agreement on the criteria for a good solution, we review our solution—our Mafia Offer. We usually give an overview of our offer and then go into each component of it in more detail. In this way, they get a preview of what’s to come—giving them the big picture—and then they can concentrate on what is being presented. The preview method also creates another *beginning*.

As we are reviewing our offer, we again deliver it in a way to which the old brain can relate. And that, of course, would change for each offer. However, it is very important that you

are *tangible*. Don't just say you have a guarantee, say what it is. Don't just say you will meet the lead time, say what it is.

For each claim or component of your offer, be very explicit about the results they will experience. Gains are typically financial, strategic, or personal. Be as *tangible*, simple, and direct as possible.

Agree Our Solution Solves Their Problem

After we have reviewed our offer, we return to the criteria for a good solution and ask if our offer has met those criteria. Then we compare our solution, our Mafia Offer or *claims*, to typical solutions to create *contrast*.

We explain that the *contrast*, the difference between us and our competitors, is what leads to the promised results in our offer and we give proof.

There are several types of proof, and here they are in preferred order: customer story or case study, a demonstration, data, or trust me.

Proof of the results your offer will provide is the core of your message. Your evidence must be *tangible*, factual, and provable. The gains you're touting must be greater than the cost of your product or service to demonstrate the value you're offering.

Close

There are many conventional close techniques out there, but if you have followed the solution for sales and TOC buy-in processes and delivered to the old reptilian brain, you don't need anything fancy. Nevertheless, we also know that the old brain pays particular attention to the *end* of a presentation. Therefore, the most effective closing

technique for the old brain is simple. Renvoise' and Morin (2007, 127-131) recommend three closing steps:

1. Repeat your offer one final time because the old brain remembers the *end*. "We will reduce your inventory by half, reducing the amount of cash you have to tie up, and at the same time we guarantee you will never stock out, reducing your chaos and costs, and allowing you to better meet your customers' needs. And if you ever do stock you out, we will pay you \$500 per day per SKU."
2. Next, go for positive public feedback by asking, "What do you think?" If you have a large group, direct this question to a particular person. Then wait for an answer. Waiting is uncomfortable but very important. The psychology of this is beautiful. The responding person will want to remain consistent with any public statement they make, and will later defend their initial position. Therefore, if they take a positive position about you or your offer, you end up with an internal advocate that remains after you leave. It's called the *Law of Consistency* (Cialdini, 2007).

It has also been found that a small initial commitment will trigger a larger commitment later (Cialdini, 2007). Have you ever noticed that *after* you purchase something you are *more sure* of the benefits you will receive? Even though just before you made the purchase decision, you were comparing and contrasting it among several alternatives? Therefore, the initial commitment that the internal advocate makes will lead to stronger statements later.

If the comments you hear are not positive, then you have the opportunity to address any concerns with everyone present. It's better to air any negatives in your presence than to have them arise when you're

not around. However, if you have done the Mafia Offer analysis well and have followed the solutions for sales process, then you will have very few, if any, objections.

3. Once you have answered all the questions and addressed any concerns, ask, "Where do we go from here?" Again, be patient and wait for an answer. Their answer is their commitment. The key to invoking the Law of Consistency is to wait *for them* to state the next steps. When your prospect finally says, "Let's pick a representative portion of labels and trial the proposed solution," it is more likely to actually happen than when you suggest it. Moreover, the person who made the suggestion will become the internal champion for the trial.

Use each presentation opportunity to improve your offer presentation and technique continuously. I find it helpful to have someone along who can help you to gauge your prospect's reaction and document any process deviations that occurred.

For Whom Can You Develop Offers?

Mafia Offers can be developed for each product or service and for each of your market segments. Some companies will have an offer for each product, while others will have offers that vary by market segment. For example, the label company uses the same offer regardless of which label is sold. However, if they decided to go into different markets, they may need a new offer for the new market. This would certainly be true if they decided to make labels for the equipment manufacturers market. In reality, they target new customers who they know suffer the effects of the price per quantity curve and an incorrect forecast.

It is common for our clients to have one or two offers for the products or services they sell in the markets in which they participate. However, there is no right number. If you need to decide which product or service and which market with which to start, you can use the same questions that we listed before:

- What market do we want to grow?
- What market has the best margins?
- Do we have too much business with one customer or in one market?
- Which customers or types of customers do we dread? (If our competitors also dread these customers, they may be more easily acquired.)
- What market has tons of room for us to grow?

Mafia Offers can be aimed at others besides your customers and prospects. You can create Mafia Offers for your vendors, your employees, your bank, your partners or affiliates, or for whomever you choose to target.

Once you start thinking along the Mafia Offer lines, you will also find it useful to ask, “Why should anyone you interact with do business with you?” This line of thinking will help you to make sure that all the interactions you have are as good as they can be. For example, if you want to start using Twitter, my question for you is, “Why should I follow you?” If you approach Twitter¹⁸ with the answer to this question in mind, you will be more successful at getting followers. Similar comments can be made about your personal interactions as well.

There is no limit to the number of offers you can develop or to the amount you can increase your sales. In addition, Mafia Offers are possible for the majority of companies. The reason that most companies don’t know that they have one

or know what it is, is because they just don't know how to develop them.

Can You Create a Mafia Offer?

If you read Dr. Goldratt's (1994) book, *It's Not Luck* or one of the other TOC books, you are familiar with how we use cause-and-effect logic (also called TOC Thinking Processes) on your customers, your industry, and your company to create the offer.

Historically, to develop a Mafia Offer you had to hire a TOC expert and spend about 2 weeks creating your Current Reality Tree, Future Reality Tree, and then your Prerequisite and Transition Trees. It is time-consuming and expensive but absolutely worth it.

As I've worked with clients, I've noticed that the majority of the population has a tough time building logic trees. They can understand them, but building them can be a challenge. Therefore, we tried another approach and have had success with a process¹⁹ that uses the logic without requiring the building of trees. So can you create a Mafia Offer? Yes, if you can delve into the logic from your customers' perspective.

In *It's Not Luck*, Dr. Goldratt develops three offers for three different companies (printing, cosmetics, and steam) in three very different environments. One was for a printing company and it was based on the price per quantity curve industry practice. Despite that, and despite the fact that all the logic is laid out, printing companies don't imitate the offer.

Even if you have trouble with the logic, you probably know enough about your industry and customers to create an offer. It's not only the logic that can be tough, but also seeing what your industry practices are. You may need an outside resource or fresh eyes to help you identify your industry practices.

Another resource that you may find helpful is the Mafia Offer templates. As with any template, there are pluses and minuses. On the plus side, templates can save time and give you some ideas about what your offer might be. On the negative side, because a template exists, you may not do the full analysis and come up short, or the offer doesn't work because it didn't fully apply and you didn't do the analysis to understand how to modify it.

We don't use templates in our Mafia Offer Boot Camps.²⁰ Instead, we do the full analysis for each company and market. It takes 2.5 days, but the mistakes are limited. So, use the templates if you must, but do the full analysis and take care to tailor them.

The Templates

Goldratt Consulting has taken the common templates and created *Strategy and Tactics Trees (S&T)*²¹ for each one. An S&T provides the roadmap to build, capitalize, and sustain a decisive competitive edge. Therefore, it includes the major operational changes that are necessary to capitalize on a Mafia Offer and it includes what needs to be done to sustain the offer and these improvements. These roadmaps can be very helpful, but only if the Mafia Offer fits for you. So don't force it, *do the analysis!*

Here is the list of S&Ts that Goldratt Consulting has published along with the Mafia Offer that goes with each one. Mafia Offers are not explicitly stated in the S&Ts, but each states the "decisive competitive edge" that will be built and capitalized upon. The decisive competitive edge is often stated with a second phase option. The second phase is an extension of the initial competitive edge, often making it even stronger. Many companies start with the initial competitive edge and work their way to the second phase.

I have added the Mafia Offer that could be made. To get the full picture, you will need to review the entire S&T, but here are the decisive competitive edges and Mafia Offers for each:

Vendor Managed Inventory

Situation: A manufacturer or distributor creates a decisive competitive advantage to make a Mafia Offer to another manufacturer or distributor.

“A decisive competitive edge is gained by providing a ‘partnership’ that guarantees remarkable availability coupled with reduced inventories and much less hassle, when all other parameters remain the same.” As a second phase, “In mature partnerships, the company has the ability to command higher prices (alternatively, to successfully defend against pressures to lower prices).”²²

This template fits the label company example we have been following throughout this chapter. Moreover, it may fit in situations where at least some of these statements are true:

- Customers/prospects are not completely satisfied with the current balance between availability and inventory.
- Repeat orders are placed for the same SKUs.
- Customers/prospects order the exact same SKU relatively infrequently.
- The value of the SKU is not negligible.
- Customers/prospects are producing and ordering essentially to a forecast.
- Life span of inventory is relatively limited—meaning the products are not good forever, but are good for a number of years.

- There are emergency orders (e.g., 3 percent).

In other words, this template may apply if your customers/prospects are holding a significant amount of inventory and, despite that, they are experiencing too many of some SKUs and stocking out of others.

Example Mafia Offer: “Mr. Customer, don’t give me orders. Your orders are based on your best guess of how many labels you think you might need. That’s because label printers put that price per quantity curve in front of you and force you to have to guess out six months. The forecast ends up being wrong, and how can it possibly be right? Instead, tell us every day how many labels you use and we can guarantee on the one hand that you won’t have to hold more than two weeks’ worth of labels. And you know how your marketing department was complaining that they can’t make the changes they want because you have six months worth of inventory? Well, you will only have two weeks. Now, at the same time we will guarantee that we never stock you out. We will guarantee that you’ll never go to the shelf and not have the label you need. And if we ever do stock you out, we will pay you \$500 per day per label. We offer all this at the same competitive price you pay today and of course you will have a lot less of your cash tied up.”

Operational Improvements Required: Implementing S-DBR²³ or the Velocity Scheduling System²⁴ and the Replenishment²⁵ solution to replenish raw materials and the customers as they consume the product.

I don’t like the name of this template because what is traditionally meant by *vendor managed inventory* (VMI) is not what is being offered here. Traditional VMI typically

requires someone to forecast and for the manufacturer to hold inventory based on that forecast. The forecast ends up being wrong, and despite the fact that the manufacturer is holding inventory, there can still be stockouts. Moreover, as manufacturers try to minimize stockouts, it's inevitable that they end up with too many of some SKUs. In this scenario, the customer still places orders. With the Mafia Offer, there are no orders. Daily consumption data is passed from the label customer to the manufacturer. The manufacturer holds zero inventory. The only inventory in the system is the two weeks' worth of inventory the label customer is holding.

Reliable Rapid Response

Situation: A manufacturer or service company creates a decisive competitive advantage to make a Mafia Offer to another manufacturer, distributor, or project-based or service company.

“A decisive competitive edge is gained by the market knowing that the company’s due-date promises are remarkably reliable, when all other parameters remain the same.” And, as a second phase, “On a considerable portion of the sales, high premiums are gained by the market knowing that the company can deliver in surprisingly short lead time.”²⁶

This template may fit in situations where at least some of these statements are true:

- The standard lead time in the industry is relatively long (e.g., ~6 weeks).
- The standard DDP in the industry is relatively poor (e.g., ~80 percent DDP).
- Customers/prospects are ordering essentially to a forecast.

- Unavailability has significant consequences for the customers/prospects.
- Customers/prospects do not find it easy to pursue an alternative solution when they are out of stock. In other words, the product is not a commodity and not readily available in the market. In addition, there is no alternative product that can be easily adjusted or modified.
- The product is highly customized and not typically sold again or the customer/prospect purchases a high number of SKUs.
- The purchase price is negligible relative to the selling price (e.g., ~5 percent, for the second phase to apply).

In other words, this template may apply if your customers/prospects suffer from unavailability, from not having your product or service (and can afford to pay a premium to eliminate the damage for the second phase).

Example Mafia Offer: “Mister Customer, we know that everyone quotes a 4-week lead time but rarely does anyone ever deliver in 4 weeks. This causes you to juggle your schedule or sometimes for your lines to go down. So, we are going to give a 4-week lead time at our current competitive pricing, but we are going to back it up with a penalty. For each day that we ship late, we will deduct 10 percent per day late off your order. And if we are 10 days late (which presently happens all the time), your order is free. [Second Phase] In addition, we know that sometimes your needs change because your customer has made changes, so we can also offer a 2-week lead time for a 2X price, but if we ship a day late we deduct 50 percent per day. And, in the rare case that you need it in 1 week, we will do whatever it takes.

This is a 4X price, but if we ship late your order will be free.”

Operational Improvements Required: Implementing S-DBR²⁷ or the Velocity Scheduling System.²⁸

Consumer Goods

Situation: A manufacturer or distributor creates a *decisive competitive advantage* to make a Mafia Offer to a consumer goods retailer or e-commerce store (that holds inventory).

“A decisive competitive edge is gained by providing a ‘partnership’ that delivers superior inventory turns (better availability coupled with substantially reduced inventories), when all other parameters remain the same.” And, as a second phase, “A decisive competitive edge is gained by providing a partnership that secures the clients an increase in TPS [Throughput Per Shelf]²⁹ and provides a realistic chance of sharing (the increase in revenues) in a much higher increase.”³⁰

This template may fit in situations where at least some of these statements are true:

- The retailer frequently “sells through” or stocks out of fast movers.
- The retailer has limited shelf space (or storage space).
- The retailer places orders with the distributor/manufacturer based on a forecast.
- A large portion of the retailer’s shelf space is taken up with relatively large quantities of slower movers.
- Slow movers are discounted after a while.
- Expecting to find an SKU and being disappointed severely erodes the consumer’s (the retailer’s

customers/prospects) impression and increases consumer disappointment.

- A long replenishment time (relatively) causes shortages and high inventories that block the shelf space and impair the ability to adjust the offering to the actual market preferences.

In other words, this template may apply if a retailer is experiencing slow inventory turns and margin erosion due to discounting of slow movers.

Example Mafia Offer: “Mister Customer, we know that everyone promises sell through and high gross margin, but places all of the risk on you to forecast and manage the inventory. If the forecast is wrong, you miss an opportunity with fast movers and then end up discounting the slow movers. So our offer is to manage our inventory on your shelf and we guarantee we will meet or exceed your historical return on shelf space or we will pay the difference.”

Operational Improvements Required: Implementing S-DBR³¹ or the Velocity Scheduling System³² (if manufacturer) and the Replenishment³³ solution to replenish consumer goods as they are sold. Switching to a mode of operations that is based on actual consumption ensures very high availability coupled with surprisingly high inventory turns and will minimize slow movers and the need for discounting.

Projects

Situation: A company that delivers projects will create this decisive competitive edge and make some version of the Mafia Offer that follows.

“A decisive competitive edge is gained by the market knowing that the company’s promises are remarkably reliable, when all other parameters remain the same. In the multi-projects arena, remarkably reliable (very high DDP without compromising on the content) is defined as delivering well over 95 percent on (or before) promised due date, while in cases of late delivery the delay is much smaller than the prevailing delays in the industry.” And, as a second phase, “On a considerable portion of the projects, bonuses (for early delivery) are gained.”³⁴

This template may fit in situations where at least some of these statements are true:

- A delay in delivery is very likely to cause a delay in the completion of the overall project.
- The standard DDP in the industry is notoriously poor.
- Late delivery of the overall project has major consequences for the client.
- The benefits of early delivery are significant and customers/prospects can afford to pay a premium to gain benefits of earlier deliveries for second phase. Moreover, customers may have even asked for or impose late delivery penalties.

In other words, this template may apply if customers/prospects frequently suffer from late deliveries and may gain substantial benefits from earlier deliveries.

Example Mafia Offer: “Mister Customer, we know everyone quotes an aggressive project lead time in an attempt to gain your business, but rarely does anyone actually complete the project in that time. These delays have a significant cost to you and delay your income. So, our quotes will include a guarantee. We will pay you

5 percent of the total project fees for each day we deliver late. At the same time, we can allow you some flexibility in making final project spec changes. In addition, we will make every attempt to deliver our project ahead of schedule, allowing you to generate income sooner, if possible. And if we can do this, we ask for a 5 percent bonus for each week we deliver early.”

Operational Improvements Required: *Critical Chain Project Management* (CCPM)³⁵ or Project Velocity System³⁶ can bring projects to have DDP of almost 100 percent. Moreover, a mature implementation can cut project lead times to be as short as 50 percent of regular lead times.

Pay Per Click

Situation: This template applies to manufacturers of equipment or capital-intensive products.

“The company gains a decisive competitive edge in large markets by providing its equipment in a way that does not involve (almost) any risk for the client.”³⁷

This template may fit in situations where at least some of these statements are true:

- The initial investment that is required to purchase the equipment is not negligible.
- The level of usage that the customer/prospect needs is highly uncertain.
- Using the equipment is beneficial to the customer/prospect.
- The initial investment required is very high for the customer/prospect.

- The income stream is tied directly to the equipment and is unstable.

In other words, this template may apply if customers/prospects want the equipment but regard the investment in the equipment as too risky. Alternatively, lack of experience causes the potential customer to doubt both the benefit and the level of usage.

Example Mafia Offer: “Mister Customer, most equipment suppliers put all the risk of purchase on you. At best, you might be offered a lease or rental agreement. We can reduce your risk to 5 percent of the purchase price and you pay only as you use the equipment. Since you pay per use, our incentive is to maximize your uptime and quality. Then you can focus on your business instead of worrying about when and if to buy.”

Operational Improvements Required: S-DBR³⁸ or the Velocity Scheduling System,³⁹ Project Velocity System,⁴⁰ or CCPM,⁴¹ depending on the situation. Using these logistical applications of TOC ensures that the current deliveries are not deteriorating and exposes excess capacity. This ensures that the only investment that is required for Pay-Per-Click offers are the *totally variable costs*⁴² of the equipment.

Gain Sharing (My Mafia Offer)

Situation: This template applies to any company supplying a product or service to any other company or individual. This template does not yet have a corresponding S&T.

This template may fit in situations where at least some of these statements are true:

- The selling company is putting all the risk on the customer/prospect.
- The customer/prospect has some doubt about the results or benefits that will occur.
- The selling company is confident in the results that can be achieved by the customer.
- The results gained by the customer can be measured.
- The customer/prospect cannot afford to purchase the product/service outright, but could pay if the promised results were realized.

In other words, this template may apply if customers/prospects want the product/service but regard the investment as too risky. Alternatively, past experience or lack of experience causes the potential customer to doubt that the promised results actually will be realized.

Example Mafia Offer: “Mister Customer, most consultants charge by day or project, putting the risk of actually getting bottom-line results on you. At best, some consultants will offer to get paid as deliverables are met, but these deliverables are typically based on the completion of some task, not your bottom line. So, our offer is that you only pay us if and when your profits increase. If we don’t increase your profits, you don’t pay.”

Operational Improvements Required: S-DBR⁴³ or the Velocity Scheduling System,⁴⁴ Replenishment,⁴⁵ Project Velocity System,⁴⁶ CCPM,⁴⁷ or nothing, depending on the situation.

Mafia Offers have been created for all different types of companies. The majority of companies, we estimate 75 to

80 percent, can develop a good Mafia Offer. Moreover, about 30 percent of the companies that have completed Mafia Offer Boot Camps⁴⁸ have been service companies. The toughest situations are:

- E-commerce sites that don't hold inventory and sell the exact same SKUs as the competition. However, there are some good offer opportunities for these companies and the key is to either (1) make the price comparison more difficult or (2) make them choose you, all else being equal.
- Those that sell insurance or sell financial planning. Again, a much improved offer and positioning is possible, but because you don't control operations, a Mafia Offer is much more difficult.
- Generally, when regulatory rules prevent the appropriate offer, for example, it is considered unethical for Certified Public Accountants to use a gain-sharing offer.

Summary

So, a Mafia Offer can help you to increase your sales by answering, "Why should I buy from you?" And if delivered correctly, using the psychology of Mafia Offers, you can have better control over your sales. What's better control? How about closing as much as 80 percent⁴⁹ of your opportunities?

Do the analysis; the worst case is you develop a better offer and a better market position. The best case is you develop a Mafia Offer, one that your customers can't refuse and your competition can't or won't match.

A Mafia Offer is only one part of how Theory of Constraints can help you to improve your marketing. However, it is a very important part and a great place to start.

References

Cialdini, R. B. 2007. *Influence: The Psychology of Persuasion*. New York: HarperCollins, [Chapter 3](#).

Damasio, A. 1995. *Descartes' Error*. New York: Harper Perennial.

Goldratt, E. M. 1994. *It's Not Luck*. Great Barrington, MA: The North River Press.

Goldratt, E. M. 1999. Satellite Program Session 5: Marketing. *Goldratt Satellite Program*. Amsterdam, The Netherlands: AYGI Limited.

Goldratt, E. M. 2008. *The Choice*. Great Barrington, MA: North River Press.

Goldratt, E. M. and Goldratt, R. 2003. The Solution for Sales. Paper presented at the 2003 TOC Upgrade Workshop, February 21-23, Cambridge, England.

Hart, L. 1975. *How the Brain Works*. New York: Basic Books.

Kim, C. and Mauborgne, R. 2005. *Blue Ocean Strategy: How to Create Uncontested Market Space and Make Competition Irrelevant*. Cambridge, MA: Harvard Business School Press.

Mabin, V. and Balderstone, S. 2000. *The World of the Theory of Constraints: A Review of the International Literature*. Boca Raton, FL: St. Lucie Press.

Ornstein, R. 1992. *The Evolution of Consciousness*. New York: Simon and Schuster.

Renvoise', P. and Morin, C. 2007. *Neuromarketing: Understanding the "Buy Buttons" in Your Customer's Brain*. Nashville, TN: Thomas Nelson.

Smith, J. 2006. *Creating Competitive Advantage*. Delhi, India: East West Books (Madras).

About the Author

Dr. Lisa Lang is considered the foremost expert in the world in applying TOC to Marketing. She is currently the President of the Science of Business and has recently served as the Global Marketing Director for Goldratt Consulting.

Dr. Lisa has a PhD in Engineering and is a TOCICO certified expert in TOC. Dr. Lisa is currently serving on the TOCICO Board of Directors.

Science of Business specializes in increasing profits of highly custom businesses and applying TOC, Lean, and Six Sigma to sales and marketing having developed the Mafia Offer Boot Camp, Velocity Scheduling System, Project Velocity System and Sales Velocity System.

Before becoming a consultant, Dr. Lisa was in operations, strategic planning, purchasing, R&D, and quality while working for Clorox, Anheuser-Busch, and Coors Brewing.

In addition to consulting, Dr. Lisa is a highly sought after Vistage/TEC speaker on "Maximizing Profitability." Dr. Lisa also provides professional keynote speeches and workshops for organizations like TLMI, ASC, NTMA, GPI, and NAPM, and

private events for corporations like TESSCO, Bostik, GE, Pfizer, Arcelor Mittal, Corus Group and Sandvik Coromant.

SECTION VI

Thinking Processes

[CHAPTER 23](#) [The TOC Thinking Processes: Their Nature and Use—Reflections and Consolidation](#)

[CHAPTER 24](#) [Daily Management with TOC](#)

[CHAPTER 25](#) [Thinking Processes including S&T Trees](#)

[CHAPTER 26](#) [Theory of Constraints for Education \(TOCfE\)](#)

[CHAPTER 27](#) [Theory of Constraints in Prisons](#)

Knowing how to think is of major importance to most of us. But how well do we think? Do we really have insightful and disciplined ways to analyze problems in either our personal lives or in organizations? This is a question for most of us; for managers, first-line supervisors and workers, for students, and others.

In this section we present strong tools for simple, logical, and focused reasoning. They include logical constructs to aid us in getting at the truth in the existing reality; its undesirable effects, core problems, and conflicts. The tools offer tests of reasoning to help assure validity of analysis. They include tools to facilitate the identification of underlying core problems, the construction of win-win solutions, and the planning of action steps to bring about necessary changes. The tools include capabilities for identifying potential negative consequences of planned

actions, negative consequences which if not seen and addressed could lead to the failure of a plan for improvement. Solutions require action for change. What actions and when? Tools for mapping How to Change, the obstacles to be overcome and the action steps for implementation are covered. Elements of cause-and-effect logic, techniques for logic diagramming, tests of logic and conflict resolution tools to help assure the integrity of solution are addressed.

The Thinking Processes are producing results in a wide range of organizations. Chapters in this section address applications in research, in education, and even in prisons. The Thinking Processes are simple enough to be used effectively by prekindergarteners but robust enough to be used on the most complex organization problems.

CHAPTER 23

The TOC Thinking Processes Their Nature and Use—Reflections and Consolidation

Victoria J. Mabin and John Davies

Nothing is more practical than a good theory¹

Introduction

Preface to the Chapter

The previous chapters have described *Theory of Constraints* (TOC) applications in various functional areas and activities, such as projects, production, accounting, strategy, sales, and marketing. All the innovations put forward in those chapters are underpinned by the powerful Thinking Processes used by Goldratt to develop solutions for common problematic situations such as those encountered in *The Goal* (Goldratt and Cox, 1984). These thinking processes were then formalized into a suite of *Thinking Processes* (TP) by Goldratt and colleagues in the early 1990s (Goldratt, 1990a, 1990b; Scheinkopf, 1999) leading to their public unveiling in *It's Not Luck* (Goldratt, 1994). As Watson et al. (2007) explain, in keeping with Goldratt's preference for the Socratic Method and directed at self-discovery, *It's Not Luck* is not a cookbook for implementation of generic TOC solutions; rather it presents a roadmap for discovering novel

solutions to complex unstructured problems. The TP provide a rigorous and systematic means to address identification and resolution of unstructured business problems related to management policies (Schragenheim and Dettmer, 2001). The TP have subsequently been described, used, and developed further by many TOC practitioners, academics, consultants, and authors.

This chapter introduces the TP, while the following chapters will describe the TP in more detail and demonstrate their use in day-to-day operations, strategic and tactical planning, and in various domains such as schools and prisons. While such applications provide many concrete and convincing examples of how the TP can liberate our thinking and change lives, they are far from exhaustive: The TP are equally applicable in every area of our lives, and are fully deserving of serious study in terms of such applicability and utility.

Copyright © 2010 by Victoria J. Mabin and John Davies.

Purpose of the Chapter

Our first aim is to provide an overview of the TP that is not only conceptual and methodological in orientation, but which also has a practical dimension based on the literature. In doing so, we seek to provide a supporting rationale for the existence of the TP by explaining how they fill a need of a methodological and practical nature not addressed by other problem-solving methods.

Our second aim is to respond to calls for more rigorous academic research on TOC, applying academic methodologies and concepts to TOC, to confirm and improve its methods, and to apply academic rigor to such research on TOC (Ronen, 2005). To this end, we note the need to review existing research on TOC TP in terms of their

methodological and theoretical basis, and examine the underpinnings of TOC from a methodological viewpoint, in the belief that by so doing, we may assist TOC in achieving its deserved recognition as a “proper” methodology. Practitioners and academics alike will ultimately benefit from such analysis.

Outline of the Chapter

In this chapter, we first provide a brief description of the nature, development, and use of the TP before examining how they relate to one another and to other typical approaches to problem structuring and problem solving. In order to make this comparison, we use extant conceptual taxonomies, not only to examine how the TP contribute to different phases of problem-solving activity, but also to examine the implicit assumptions and underlying philosophical frameworks that characterize TOC and other approaches. This allows us to better understand TOC as a methodological set, its strengths, and its potential for development, using TOC TP tools and methods on their own, in concert with one another, or with other decision-making methods. In doing so, we provide an alternative route or means by which to validate and enhance the TP.

The investigation and identification of TOC as a methodology or meta-methodology also allows us to see TOC as more than just a set of problem-solving logic tools: TOC fits well with a philosophy of continuous improvement, as well as prompting dramatic change; it fits with other systems mapping methods and problem-solving approaches such as *Operations Research/Management Science* (OR/MS—both *hard* and *soft* OR). The examination of TOC’s philosophical underpinnings and the comparison with other methodologies provides a basis for TOC to be viewed as a legitimate field for academic enquiry, not just as a problem-

solving methodology. When used as problem-solving methods and tools, the TP allow managers to draw on the relationships between causes and effects, between end goals and their necessary conditions, to build pictures of their realities, capturing complexity, viewing conflicts, yet still being able to discern a way forward. The tools handle complexity and systemic interactions without losing sight of the key factors: the core problems and thorny dilemmas that need resolving to make true progress.

The Nature, Development, and Use of the TOC TP

In this section, we provide a brief overview of the TP, and of their historical development from their early published forms to the current day. We comment on their underpinning logics, and describe the TP in order to discuss the categorization of the TP literature that highlight the use of the TP, and then in the following section, to explore the philosophical and methodological characteristics of the TP. Such categorization then facilitates a deeper understanding of why the TP are considered to be systemic in nature, and why the TP have been termed a “complete package” by Dettmer ([Chapter 19](#), this volume) or a comprehensive methodology or meta-methodology by Davies et al., (2005).

Overview of TP and Their History and Development

Watson et al. (2007) provided a “Silver Anniversary” review of the evolution of TOC concepts and practices, reviewing TOC’s accomplishments and deficiencies. The development of the TOC approach began with a manufacturing scheduling algorithm in 1979, which tripled plant output in a short time, and was reported at a 1980 APICS conference. Its development continued as an effective methodology for

production applications (Cox and Spencer, 1998), and by the mid-1990s, the approach was in worldwide use by companies of all sizes (Hrisak, 1995). Goldratt (1994) then developed a suite of logic tools to help managers address business problems in general. These have become known collectively as the TOC Thinking Processes, the TP logic tools, or TP tools (see Kendall, 1998; Dettmer, 1998; Scheinkopf, 1999), although Dettmer chooses to use the term *Logical Thinking Process* (LTP) to describe a modified and expanded set of thinking processes that have been developed to address issues of a strategic nature (Dettmer, 2007; [Chapter 19](#) in this Handbook).

The TP tools act as guides for the decision-making process as well as representations of logic. They embrace problem structuring or representational tools, such as the *Current Reality Tree* (CRT), the *Evaporating Cloud* (EC), and the *Future Reality Tree* (FRT), and tools such as the *Prerequisite Tree* (PRT) and *Transition Tree* (TRT) that facilitate effective implementation.

The TP were developed to facilitate beneficial change, which in most circumstances also requires, or relates to, overcoming resistance to change. They guide the user to find answers to basic questions relating to the change sequence, namely, *What to Change? What to Change to?* and *How to Cause the Change?* For example, the CRT helps identify what, in the system, needs to be changed. The EC is then used to gain an understanding of the conflict within the system environment, or of the reality that may be causing the conflict. The EC also provides ideas of what can be changed to break the conflict and resolve the core problem. The FRT used in concert with the *Negative Branch Reservation* (NBR; a sub-tree of the FRT) takes these ideas for change and demonstrates that the new reality created would lead, in fact, to resolution of the unsatisfactory systems conditions and not cause new ones. The PRT determines obstacles to implementation and the desired

sequence to overcome them, and the TRT is a means by which to create a step-by-step change implementation plan. Four preliminary steps usually precede such discussion, namely, *What the system is*, *What its goal is*, *How progress toward the goal will be measured*,² and *Why the change is needed*. In addition, following these are the steps to sustain the change and to develop a *process of ongoing improvement* (POOGI).³ Dettmer (2007) provides the *Intermediate Objectives* (IO) map for this purpose, while others follow the Business System Model of Cox et al. (2003) and the Three-Cloud Method (Button, 1999, 2000), and yet others describe these steps as preliminary to the *Five Focusing Steps* (5FS; Scheinkopf, 1999).

The various TP tools subsequently have been further developed to improve or simplify the building of logic diagrams. While the TP were designed and introduced as an integrated set of problem-solving tools, we know also that by using the TP tools, individually or in concert, an organization can develop and implement change solutions successfully (Scheinkopf, 1999).

The TP's embrace and are constructed from three basic logic building blocks (Scheinkopf, 1999). Two of the building blocks manifest cause-effect thinking through employing either sufficiency-based *if-then* logic or necessity-based *in order to . . . we must have . . .* logic. The CRT, FRT, and TRT are sufficiency-based logic diagrams, whereas the EC and PRT are necessity-based logic structures. The third building block manifests as a set of rules governing the logic-in-use and provides a protocol for establishing and challenging the existing cause-effect thinking and logic. It does so through the seven *Categories of Legitimate Reservation* (CLR) (Goldratt, 1994, [Chapter 15](#); Noreen et al., 1995; Dettmer, 1998; Scheinkopf, 1999, [Chapter 4](#). [Chapter 25](#), Appendix B of this Handbook lists and describes the CLR.) that legitimize, depersonalize, and depoliticize any challenges to

current thinking. Such rules are used to add rigor to the modeling process and to check the validity of the constructed logic relations as logic tree diagrams. The result is a logical, structured, and rigorous process to guide managerial decision-making, utilizing the intuition and knowledge of those involved and invoking challenges to existing thinking using the protocols of the CLR.

The next section provides a description of each TP tool, judged to be sufficient to characterize and facilitate comparison of the TP tools and methods with other more traditional OR/MS methods. Further description is provided in the following chapters, and formal definitions in Sullivan et al. (2007).

The TP Tools

The Current Reality Tree (CRT)

*The CRT is a sufficiency (if-then) logic-based tool used to identify and describe cause-effect relationships that may help to determine core problems that cause the undesirable effects (UDEs) of the system (Cox et al., 2003; Sullivan et al., 2007). The CRT is designed to answer the question, *What to Change?* taking care to avoid actions that merely deal with symptoms. This tool is a particularly effective tool if the symptoms are caused by a policy as opposed to a physical constraint of the existing system. A useful variant is the *Communications Current Reality Tree* (CCRT; Scheinkopf, 1999; [Chapter 25](#), this volume).*

The Evaporating Cloud (EC)

Policy constraints identified in the CRT can often be viewed as a conflict or dilemma between two opposing actions. The TP tool for such situations is the *Evaporating Cloud* (EC), referred to by some as the *Conflict Resolution Diagram* (CRD; Dettmer, 1999). *The EC is used for solving problems—using necessity-based (in order to, we must . . .) logic—that may arise not only from the seeming irreconcilability of opposing actions, attitudes, and behaviors, but also from what may be regarded as a chronic conflict of competing actions, conflict of interest, or as intractable dilemmas of a political, policy, or ethical nature.*

Though the EC process frames the problem, for example, as starting with two diametrically opposed actions or views, engaging in the process also implicitly assumes these matters can be resolved by a win-win solution to generate the system goal or objective A, via the attainment of necessary intermediate states, B and C. In order to find such a solution, we elicit those assumptions or reasons why the relationships are thought to hold. Some of these assumptions may be shown as annotations in the “thought bubbles” on the EC diagram ([Fig. 23-1](#)).

Often when the assumptions are surfaced and articulated, they may be seen to be false or weak, and the conflict represented by the cloud evaporates. Where assumptions are recognized as valid, they may be addressed in a manner that invalidates them, reduces their importance or impact, and allows for a resolution of the conflict. We develop a list of such assumptions and the accompanying “injections” that may be used to “attack” or address those assumptions to resolve the conflict. Indeed, the EC diagram may provide a basis for insights about the nature of root causes and the core problem identified in our CRT. Specialized versions of the EC include the *Generic Evaporating Cloud* (GEC), the 3-UDE Cloud, and the Core Conflict Cloud. See [Chapters 24](#) and [25](#) for a detailed development of the EC.

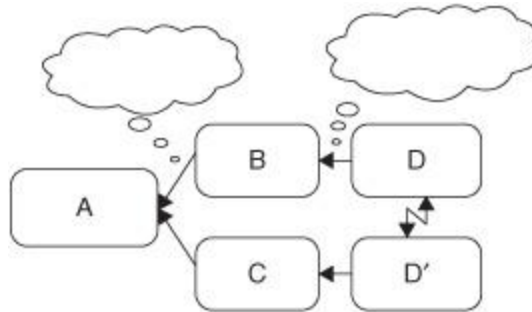


FIGURE 23-1 The EC diagram.

The Future Reality Tree (FRT)

The FRT process, in contrast to the CRT, begins with the identification of actions, conditions, or solutions of choice, what Goldratt collectively names as “injections,” and then through the mapping of sufficiency-based logic relations, checks whether the causal links will lead to what we have decided are desirable outcomes, that is, the removal or closing of Dettmer’s “mismatches.” As Rizzo (2001, 14) states, the construction of the FRT can be viewed as a “what-if exercise,” helping to identify what actions and conditions will be necessary and sufficient to bring about desirable effects or change, and whether or not additional UDEs will also emerge from our actions (Kendall, 1998, 39).

Subtrees may be constructed in this process whenever someone raises a “Yes, but . . .” doubt or type of reservation. Such situations indicate that the “objector” has thought of a possible negative side effect of the proposed solution. Rather than brush the comments aside or abandon the proposal, we are encouraged by the TOC philosophy to explore ways of adapting the proposal to avoid such negative side effects while still keeping the positive effects, using a method known as the *negative branch reservation (NBR)*. *The NBR (Goldratt, 1996) is formally a sub-tree of the FRT, but can be used as a stand-alone tool to improve*

critical feedback and develop half-formed ideas such as changes to organizational performance measures. Illustrations of the NBR method can be found in Boyd and Cox (1997), Mabin, Davies, and Cox (2006), Dettmer (2007, 226).

The Prerequisite Tree (PRT)

Development of the PRT, complementing and building on the FRT, seeks to identify local obstacles, omissions, and conditions that might block the path to the desired outcomes, and then to set new IOs and goals that would equate to overcoming those obstacles. The PRT is often developed by a team, in addressing obstacles that confront them, and hence social practices and power relations embedded in the problem will be considered implicitly, if not explicitly. If the team or working relationships are perceived to be an obstacle, then such issues will usually be raised.

The Transition Tree (TRT)

The development of the final logic structure, the TRT, seeks to identify tasks and actions both necessary and sufficient to meet the IOs of the PRT, to overcome what might go wrong, to provide a rationale and schedule for action, and, as such, to provide what we may regard as a coherent step-by-step implementation plan, and which also accounts for prevailing beliefs, feelings, and norms.

Summary

As we move through the tools, CRT through to TRT, there is generally more involvement from the wider group affected by the problem, or by actions designed to address it. The tools purposefully address successive layers of “resistance”

and “buy-in” (Houle and Burton-Houle, 1998; Goldratt, [Chapter 20](#); Lang, [Chapter 22](#)), and other issues raised in the broader “change management” literature (Mabin et al., 2001). The CRT may be developed by a smaller group, initially, with buy-in being developed increasingly through the remaining steps of the TP. Likewise, empowerment also develops through the TP. The CRT represents the current situation, enlightening but not necessarily empowering. The PRT and TRT in particular are designed to build collective buy-in, aiding the implementation phase. The end goal and normal outcome of the FRT, NBR, PRT, and TRT is to help people gain a better understanding of the problematic situation and the results of their actions, and to feel empowered through having an agreed course of action.

The next section moves from a consideration of what the TP tools are to a review of the tools-in-use, patterns of use, and opportunities for further use and enhancement of the tools.

The TOC TP Literature

In this section, we review developments to the TOC body of knowledge, particularly the TOC TP as reported in the public domain peer-reviewed literature. In doing so, we also comment on the nature of the TP, vis-à-vis their evolution and their domains of application.

The commentary primarily draws on the work of Kim et al. (2008) who examined the peer-reviewed literature on TP, from the publication of Goldratt’s *It’s Not Luck* in 1994 up until early 2006. Two prior studies, by Rahman (1998) and Mabin and Balderstone (2000; 2003), provided reviews of the broader TOC literature, and reviewed papers published before 2000. Kim et al.’s (2008) work complemented and extended those other reviews by focusing on TP up to early 2006. These reviews have provided a valuable summary, for

academics and practitioners, of the developing TOC body of knowledge that have found outlet in the peer-reviewed literature. In addition, Watson et al.'s (2007) review of the evolution of TOC, while not attempting to provide a literature review, does discuss TP and identifies some deficiencies.

Whereas Rahman's (1998) review of the TOC literature classified the TOC literature based on what he termed the philosophical orientation and application of TOC, the review conducted by Kim et al. (2008) used an extended set of five dimensions or orientations: theoretical or methodological, application, time, epistemological, and TP tool orientation. Kim et al.'s review relates to over 110 peer reviewed journal papers on TP, 70 of which were published in the period from 2000 to early 2006. A subsequent search reveals another dozen or so applications papers published between early 2006 and late 2009.⁴ We summarize and update the main findings from the Kim et al. review in the next sections, looking at application orientation, the prevalence of individual tools, and last, methodological developments.

The Application Orientation of the TOC TP Literature

Over 100 papers have described applications of TP. Kim et al. (2008) identified three self-defining categories of TP "application-oriented" papers, namely those relating to the whole business system, to specific functional areas, and to the service sector. Applications to the whole business system mainly described the process of implementing the use of TP tools in a single organization, and investigated the impact of TP on the organization in terms of organization-wide performance measurement and change management. Such applications traversed a diversity of issues and contexts including change management, performance

measures, pricing conflict, outsourcing decisions, project cost recovery, mergers, and healthcare.

TP applications to functional areas included manufacturing and production, *Supply Chain Management* (SCM) in particular, but also marketing, sales, accounting, quality, strategy, human resource management, and new venture development, addressing outdated policies, unacceptable scrap rates, and poor delivery performance. SCM applications included identifying critical success factors and a performance measurement system to assist supply chain members to realize the potential benefits of collaboration. Recent papers address invoicing (Taylor and Thomas, 2008) and human resource management functions (Taylor and Poyner, 2008).

About a third of the TP application-oriented papers described how TP have been or could be applied to service sectors such as healthcare (for example, military medical service, operating room utilization, aeromedical evacuation system, ambulatory care system, supervisory oversight procedures, multi-site medical practices, and insurance claims processes), education (including curriculum applications and capacity management in distance education), and public services (water systems [Reid and Shoemaker, 2006; Shoemaker and Reid, 2006], and police/fire services [Taylor et al., 2006]). Legal service and white-collar service TP applications were detailed in Kim et al. (2008). In addition, there have been a number of books, especially recently, devoted to the service sector, such as Ricketts (2008), Jamieson (2007), Ronen et al. (2006), and Wright and King (2006), although these are not included in the data, as noted earlier.

TP Tool Orientation

Kim et al. (2008) also categorized papers according to the TP tools that were employed to address problem situations. The updated data confirms that by far the most common tool employed was the EC, with approximately three-quarters of papers (78 percent) using this tool, one-quarter using the EC on its own (25 percent), and over half using the EC in combination with other tools (54 percent). Nearly two-thirds of the papers (65 percent) used the CRT or one of its variants. See Tables 23-1 and 23-2.

One in eight applications papers (12 percent) used the full TP analysis, whereas over 40 percent (43/106) involved only one TP tool. Even though TP have been developed to make mutual and complementary contributions as a suite of integrated logic tools, the literature suggests that individual use of single TP tools, or tools in pairs or trios, is not only possible but has been to be found very valuable in dealing with problematic situations. See [Table 23-2](#).

Methodological Developments and Enhancements

Kim et al.'s (2008) review also identified many methodological developments and variations that have emerged, including alternative approaches to building a tree, using specific TP tools to serve a different purpose from that originally intended, using TP tools to complement the use of other tools in addressing problem situations, and the development of new TP tools, such as the GEC, the CCRT, and IO maps.

Developments Pertaining to the Building and Presentation of the CRT The CCRT is a stripped-down CRT that facilitates communication with managers. It also serves to enhance buy-in by starting from a positive proposition (the desired objective, box A in the EC) rather than just a negative one (the core problem), and shows the relationship between these and the observed UDEs (Scheinkopf, 1999,

[Chapter 12](#); Houle and Burton-Houle, 1998; Button, 1999). Button also presents Goldratt's three-cloud approach for building a CRT, which was developed to reduce the time and difficulty of building a CRT. The traditional approach incorporates a 10-step procedure seeking likely causes for observed UDEs. The 3-UDE EC approach uses four steps to construct a CRT: (1) identify a list of UDEs; (2) generate three ECs (from seemingly unrelated problems) from the list of UDEs; (3) construct a GEC from the three ECs, thus identifying the likely core conflict; and (4) build a CRT that starts with the core conflict and harnesses the logic and pictorial representation of the GEC. While Dettmer (2007) decries such an approach, it is recommended in other texts (Cox et al., 2003), and the Kim et al. (2008) review reveals that both approaches have been much in use. Dettmer instead prefers the IO map as the starting point for TP analysis, arguing that the core conflict is more likely to be identified by the IO map's more strategic approach to identification of UDEs. There may also be occasions when three clouds may not lend themselves to a GEC when one EC is nested or embedded in the other EC (Davies and Mabin, 2009). The *Strategy and Tactics* (S&T) trees—covered in [Chapter 25](#)—were not evident in the review of the peer-reviewed literature, but are being used increasingly by TOC developers and practitioners.

TP Tools-in-Use (Totals)	# of Papers Reporting Use	% of N (= Papers Reporting Use)
EC/GEC	83	78
CCRT/CRT	69	65
NBR/FRT	47	44
PRT	21	20
TRT	18	17
CLR	1	1
Total # of papers = N	106	100
Total # of reported uses = n	239	

TABLE 23-1 TP Tools—Reported Usage—1994 to 2009

TP Tools-in-Use	N = # of Papers Reporting Use	% of N (= Papers Reporting Use)
CRT (on its own)	14	13
EC (on its own)	26	25
NBR (on its own)	1	1
PRT (on its own)	1	1
CLR (on its own)	1	1
CRT, EC	16	15
CRT, EC, FRT	13	12
CRT, EC, NBR	1	1
CRT, FRT, NBR, TRT	2	2
CRT, FRT, (NBR)	2	2
CRT, EC, FRT, PRT	1	1
CRT, FRT, PRT, TRT	1	1
CRT, EC, FRT, PRT, TRT	1	1
GEC, CRT, FRT, (NBR)	5	5
EC, FRT	1	1
EC, PRT	1	1
EC, FRT, PRT	2	2
EC/NBR	3	3
FRT (NBR), PRT, TRT	1	1
Full TP Analysis	13	12
Total # of papers = N	106	100%

TABLE 23-2 Classification of the Literature by TP Tools-in-Use —1994 to 2009

Methods Being Used Singly or in Sequenced Use Once TOC practitioners have identified *What to Change* by using the CRT, the second step in the traditional TP approach deals with the search for a plausible solution to the root cause; that is, *to What to Change*. This task can be accomplished with the aid of the EC and the FRT (see, for example, Taylor and Thomas, 2008; Taylor and Poyner, 2008). As evident from [Table 23-2](#), many authors have seen

the advantage of the EC as a standalone tool or method, and how it can lead to a win-win solution by surfacing and breaking the assumptions underlying the supposed conflict. The papers reviewed by Kim et al. (2008) described use of the EC method in conflict situations as varied as interpersonal conflict between sales manager and salesperson, writing MIS mini-cases, the creative design process, SCM, resource allocation in schools, TOC education, forest harvesting, Lean manufacturing and TOC implementation, managerial dilemmas, and “traditional” measurement. (See Kim et al., 2008 for further details.)

Nevertheless, it has also been suggested that the use of the EC following development of the traditional CRT provides potentially far more diagnostic and solution generation power than individual use of the EC or CRT. One reason is that once a core problem has been identified using a CRT, it is more likely that a solution can be developed using the EC. Several papers (see Moura, 1999; Smith and Pretorius, 2003; Choe and Herman, 2004; Umble et al., 2006) describe and explain the combined use of both TP tools to identify the system’s core problem and possible solution.

CRT-EC-FRT Method and EC-CRT(B)-FRT(B)-NBR Method Other variants on the “traditional” approach (CRT-EC-FRT) include the GEC-CRB-FRB-NBR multi-method approach, a refinement using the *Current Reality Branch* (CRB) and *Future Reality Branch* (FRB) (Cox et al., 2003). Cox et al. (2005) suggest using the EC rotated clockwise to provide a skeletal structure for the middle/upper part of the CRT. This and related papers (e.g., Davies and Mabin, 2007; 2009) also combine an EC portrayal of the conflict with a *causal loop diagram* (CLD; Senge, 1990) from System Dynamics (SD), which reveals the nature of systems relationship as well as identifies the underlying systemic structure of the problem situation. They have found that an important aspect of the EC process is that it ensures that

the system goal is reflected appropriately in a second modified CLD, when it may have remained implicit or been overlooked in an initial CLD (or CRT) representation.

Validation Using CLR TOC's CLR provide guidelines for communicating any doubts or concerns about the validity of the entities and their connections within TP trees (see Dettmer, 1997). Balderstone (1999) suggested using the CLR for validating *System Dynamics* (SD) models, while Koljonen and Reid (1999) demonstrate use of SD models to validate TOC logic trees.

Full Thinking Processes Analysis (FTPA) While the TP logic tools or trees and the EC were developed as a suite, the review of the literature conducted by Kim et al. (2008) highlighted the frequent reported use of individual TP and the application of TP for a different purpose from that for which they had been originally designed. The latter uses, however, in no way deny the effectiveness of the TP tools as a suite that contributes to a *Full Thinking Processes Analysis* (FTPA). In a later section, we suggest why the FTPA may be regarded as a complete “package” or comprehensive methodology. As such, as designed, the FTPA would use all five original TP tools to examine a system in order to identify the core problem, develop solutions, and determine the implementation steps.

Nevertheless, the literature shows that the FTPA is often used and has value in seeking to overcome resistance to change by creating a logic path that can be followed by all stakeholders and participants. Houle and Burton-Houle (1998) lay out five layers of resistance, and correspondingly five phases of buy-in. Foster (2001) discussed nine layers of resistance to change and suggested that TP tools can be used to overcome each layer of resistance. Mabin et al. (2001) relate the layers of resistance to the sources of resistance identified in the change management literature and link the TP tools accordingly. [Table 23-2](#) shows that only 13 papers in the published literature surveyed have

contained complete descriptions of use of the FTPA—perhaps because the length of these analyses may prohibit the acceptance of such research in most journals. These papers detail how the FTPA can be applied to specific business situations (Klein and DeBruine, 1995; Boyd et al., 2001; Mabin et al., 2001; Reid et al., 2002; Gupta et al., 2004; Ritson and Waterfield, 2005; Reid and Shoemaker, 2006; Shoemaker and Reid, 2006), with other authors discussing the possibility of multi-methodology in detail (Thompson, 2003; Davies et al., 2005; Schragenheim and Passal, 2005). However, the reports of FTPA use demonstrate its versatility and applicability in relation to different functionality and settings, including establishing management policies, strategic planning, executing a bank merger, and in industry settings as different as the manufacturing industry, the motion picture industry, and the healthcare service sector.

The literature appears to support the views of TP developers, including Goldratt (1994), Scheinkopf (1999), and Dettmer (1997), that each TP tool in the TP set is a potentially valuable tool in its own right, without regard to its contribution in a suite or sequenced use of tools.

Summary of the Literature Review

The development of the TOC body of knowledge has been largely practice-led, manifested not only in the diverse nature of applications areas and in the diverse use of TOC tools, but also in the broader evolution of TOC methodology, methods, and tools. While the TOC TP had their origin and arose from concepts developed primarily in operations management, we note how their contribution to the development of the TOC body of knowledge has since generated impact well beyond the particular domain of

operations management not just to wider business, but also organizations in general.

Earlier reviews of the literature (Rahman, 1998; Mabin and Balderstone, 2003) preceded many of the developments documented here, which have evolved since 2000. This overview has drawn primarily on the work of Kim et al. (2008) to present a review of the TP literature, as published in refereed journals and conference proceedings over a 16-year period from 1994 to late 2009, and to portray the development of TP concepts and tools since first applied in the POM (production/operations management) and OR/MS domains. The review of Kim et al. (2008) revealed specific publication and research gaps, and some common future research topics and approaches have also been identified. These will be discussed in the final section.

The review of TP tools-in-use has found that a combination of tools is often applied pragmatically according to the problem situation. Indeed, the overview has positioned the many TOC tools in multi-methodological use and in relation to each other, as well as capturing developments in multi-methodological usage across several domains. Consequently, a later section will examine the design-for-purpose and philosophical basis of the TP tools, as a means of understanding whether use of a TP tool for an alternative purpose is appropriate, and whether and how the TP tools in combination or as a suite comprise a comprehensive multi-methodological set. As a corollary, we will develop alternative perspectives on the nature of TOC methods and the TOC TP tools, their philosophical basis, and their use in problem-solving activities, that will facilitate comparison with other problem structuring and problem-solving methodologies and provide insight about the communality and complementarity of such approaches and methodologies.

It is apparent that TP have become a problem-solving method of choice for many, on their own, and sometimes in

combination with other methods. Before we investigate TOC's philosophical roots, we will briefly discuss other managerial problem-solving methods in order to make a comparison.

The Nature of Other Approaches to Problem-Solving and Decision Making

The purpose of this section is to establish what other methods are being used for problem solving, in what ways they are being used, and in what ways they may be limited, thus providing a partial justification for TOC TP as an alternative or complementary approach.

As a facilitating framework for this discussion, we draw on the work of Mingers and Brocklesby (M-B) (1997) to clarify the role, function, and purpose of different problem-solving methods or tools, and for relating those methods or tools to problem content and problem-solving activity. In doing so, we seek to provide a basis for some selective comparison of traditional methods and TOC methodology.

The Relationship of Problem-Solving Methods to Problem-Solving Activity

M-B developed a two-dimensional mapping grid (see [Table 23-3](#)) with the purpose of alerting practitioners to the appropriateness of using different methodologies in different contexts, and to the possible use of multi-methodology. One dimension relates to the problem domain, specifically the nature of the world—social, personal, or material—being investigated, and a second relates to aspects of methodology, particularly the conceptually distinct but related phases of “intervention.”

These phases are described within the M-B framework, for example, as building an *appreciation* of the social, personal,

or material world that provides a necessary base for *analysis* of that world and relationships between key entities, before developing and *assessing* alternative futures and options to bring them about, and then finally being able to choose and *implement* alternative courses of *action* that bring about the desired future.

Despite Mingers' (2003, 560) later reservations about the limitations of the two-dimensional M-B framework in seeking to link methodology and method to problem content and problem-solving activity, we see value in using the M-B framework, both on its own, and in tandem with Mingers' later classificatory framework (2003).

Unstructured Approaches—Management on the Hoof

Management books paint a gloomy picture of the problem-solving and decision-making abilities of managers and organizational decision makers (Simon et al., 1987), highlighting the decision traps faced by managers (Russo and Schoemaker, 1989) and the common failings of managers (Nutt, 2002). These include, for example, weaknesses in the *appreciation, analysis, assessment, and action* phases of problem intervention—a failure to appropriately frame decision problems or problem situations; a failure in direction setting—that is, to determine inclusive, acceptable strategic goals and values; a tendency to jump in and act precipitously; a failure to understand or accommodate stakeholder influences and needs; a tendency toward overconfidence and to overestimate one's predictive ability, sphere of influence, or influence on past successes and future outcomes; a failure to learn from prior actions; a failure to recognize or address ethical dilemmas or the importance of ethical values, etc. (Russo and Schoemaker, 1989; Senge, 1990; Bazerman, 1996; Nutt, 2002). Some consequences are what many

perceive to be the predominance of a firefighting mentality, and the preponderant use of managerial fads and fashions, such as quality circles, JIT, BPR, Six Sigma, etc.—with managers having the expectation that the use of these tools or processes, even in isolation, will help address their wider problems and deliver riches, now and in the future.

Phases Dimensions	Sensibility: Awareness, Empathy, Appreciation of . . .	Analysis Understanding and Synthesis of . . .	Appraisal, Evaluation, Assessment of . . .	Purposeful Action Choices to . . .
Personal	Individuals' ideas, beliefs, meanings, emotions, aims, needs and wants	Different perspectives, perceptions and worldviews— <i>Weltanschauung</i>	Alternative conceptualizations and constructions of reality	Create common ground, and consensus about ideas, states, etc.
Social	Social context, norms, practices, relationships, power relations	Misperceptions, misrepresentations, distortions, conflicts of interests	Alternative means of critiquing, contesting, or modifying power relationships	Generate understanding and empowerment to effect desired relationships, states, etc.
Material	Physical context and relationships	Underlying causal relationships and structure	Alternative physical and structural arrangements	Identify, select, and implement best alternatives

Adapted from Mingers and Brocklesby (1997).

TABLE 23-3 Framework for Mapping Methodologies

However, if and when “managing becomes a constant juggling act of deciding where to allocate overworked people and which incipient crisis to ignore for the moment” (Bohn, 2000, 83), it is usually deemed to be more expedient to attend to the squeaky wheel, to search for local solutions, as the big picture drifts out of sight. Managers then face the issues of tackling problems as they resurface or have adverse impact elsewhere. They may have framed their problems inappropriately, tackled the wrong problems,

attacked problems at the wrong levels, or just addressed them in poor fashion.

Problems poorly addressed create more problems and take longer to fix in the long-term. Senge (1990) describes this common behavior in his Fixes that Fail and Shifting the Burden archetypes. In Fixes that Fail, an inappropriate fix might work in the short term but make the problem worse in the long term; for example, smoking may bring short-term relief but leads to long-term addiction and health problems. In the Shifting the Burden archetype, the quick fix not only makes the problem worse in the long term, but also undermines the effectiveness of any other alternative fix that could be used. For example, employing consultants may assist in the short term, but may consume resources required to develop expertise in-house.

In all of these situations, several features usually stand out. They include the lack of an overall perspective, the systems or holistic view; and a related inability to think about the wider stakeholder community, their values and views, about the wider systemic consequences over time, that is, within and without the system. More specifically, they include the related inability to think about the time-related dynamic nature of cause-effect relations and feedback. They include behavior of seeming irrationality and behavior suggesting lack of awareness of the values and perspectives of others. In essence, these are features suggesting that some formal processes may be needed—in particular, processes adopting a systems perspective. We provide a brief overview of such processes or approaches in the following subsections.

Formal or Structured Approaches

There are many structured approaches, and we have chosen to review a few of the approaches that have been compared

with TOC. In the next section, we briefly outline some of the major traditional or “hard” OR/MS approaches before providing an overview of “soft” approaches, in order to provide a comparative critique of hard, soft, and TOC methods.

OR/MS Structured Approaches

OR/MS has adopted the phrase, *The Science of Better*, describing itself as the scientific approach to solving business problems. Similar terms have been used to describe the TP. Hence, a comparison between OR/MS and TOC would appear to be appropriate. Despite its origins as a problem-focused, multidisciplinary activity employing top scientists to attack operational problems, OR/MS has become very focused on techniques. In the United States, these techniques are almost exclusively quantitative in nature, and most modern-day OR/MS textbooks are rooted in the language of mathematics: mathematical modeling in its various forms such as mathematical programming (including linear and integer programming), simulation, heuristics, scheduling, decision analysis, data envelopment analysis, inventory control, and project scheduling. In these areas, OR/MS has achieved notable successes, largely through the use of powerful mathematical and computer modeling techniques to crack large problems. As such, OR/MS tools and techniques have predominantly contributed to the *analysis* and *assessment* phases of the problem intervention process, set out in the M-B framework. Indeed, such emphasis on mathematics and its use in this way is well recognized and even reinforced by the publication regimes of the top American OR/MS journals, which restrict their scope to those papers containing mathematically rigorous treatment (Simchi-Levi, 2009).

Some leading OR/MS authors, however, view this narrow definition of OR/MS—a collection of powerful mathematical tools—as unhelpful, even detrimental to achieving the full potential of OR/MS. As Daellenbach (1994, 112) puts it:

When reading about how to do a problem formulation, the tyro management scientist is often somewhat impatient: “This seems to be all obvious—let’s get down to the really interesting mathematical modelling phase! That is real OR/MS!” Unfortunately unless the groundwork for the modelling phase is properly done in the formulation, the risk is great that, although challenging, the modelling may address the wrong problem. Not only can this have serious consequences for the analyst, it also puts OR/MS into disrepute.

Refreshingly, perhaps pointedly, Daellenbach devotes the early chapters of his text to systems thinking, systems concepts, systems modeling, and problem formulation before introducing mathematical modeling.

Many OR/MS writers have made similar points about the tendency to solve the wrong problem; for example, Gass (1989), Zeleny (1981), Rosenhead (1989), and Mabin and Gibson (1998), and offered alternatives (e.g. Pidd, 1996). The debate that raged in OR/MS circles in the 1970s, led by Ackoff (1977; 1978; 1979), was largely due to this concern that the obsession OR/MS had with mathematical modeling led the OR profession astray. TOC writers have added their voices: Jackson et al. (1994) provided a powerful case comparing the standard OR-derived *Economic Order Quantity* (EOQ) with the EC approach for inventory control, following Goldratt’s own treatment of batch sizing decisions (Goldratt, 1990b, 43); Mabin et al. (2009) compared OR’s math programming approach with an EC approach to a warehouse/distribution problem. The concern of these authors over problem definition—rather than merely

problem solution—is shared by the developers of the various *soft OR* methods, also known as *Problem-Structuring Methods* (PSMs), which are discussed in the next section.

Present-day OR/MS tools and methods have much to offer in addressing complexities relating to scale, time, and computation. They have much to offer when the problem is well defined and when goals, local or global, are known, understood, and accepted by stakeholders with common perspectives; when desired outcomes can be guaranteed by action; when the successful accommodation of multiple objectives is unambiguous; and objectives can be quantified. However, even when these conditions are not met, the sophistication of analysis and the scale of computer power can generate a sense of false security especially where the relationships of local goals to systems goals are not understood or accounted for, or when local goals surface as “numerical” or binding constraints in the hard mathematical formulations—and do so without being questioned. Indeed, even when problems are not well defined, assumptions are often made to make the problem tractable or amenable to mathematical formulation, often without sufficient questioning of the appropriateness of such assumptions.

The value of the TOC TP as a comprehensive methodology is that they bring such issues to the fore, forcing a consideration of the broader problem situation, global and local goals, challenging the assumptions that underpin them, oftentimes setting the solution path toward a very different goal.

While it is often claimed that effective OR/MS practitioners do seek to achieve the global and systems goals, the reality is often a suboptimization of a technical subsystem (in the *material* world of the M-B grid) that can be modeled, undertaken without any mandate or ability to place the problem in a wider context or to consider broader issues and ramifications.

In brief, most OR/MS methods have strengths in evaluating the relative effectiveness of alternative choices and decisions, and identifying the best among them, according to prescribed quantitative criteria and objectives, or in some cases, from a prescribed or readily imputed list of choices (for example, optimization by linear programming (LP)). These methods stop short of guiding decisions on value systems, strategic direction, or other matters of identifying strategic choice for a variety of stakeholders. Furthermore, if methods of constrained or numerical optimization are examples of such methods, then by contrast, it is the methods of *soft* OR, along with TOC, that have been designed to grapple with the *wicked problems* or *messes* that are beyond the scope of the traditional mathematical modeling methods of OR/MS (Mingers, 2009a). We explore these matters further in the next section.

Soft OR

The concern of some OR writers over problem identification and problem definition—rather than merely problem solution—is shared by the developers of various *soft OR* methods or PSMs. These were showcased in the book *Rational Analysis for a Problematic World* (Rosenhead, 1989), which became the most referenced book in the *Journal of the Operational Research Society* in the next decade (Rosenhead, 2009). *Soft OR* is suited to messy situations, where the first issue is that of *not* knowing what the problem is. *Soft OR* methods have been designed and developed to grapple with *wicked problems* or *messes* by seeking to gain understanding of what would be desirable and appropriate goals of the organizational system and subsystems, and by seeking a broader, often predominantly

qualitative, understanding of the problem domain or wider system within which it sits.

Soft OR or PSMs aim to:

- structure complexity of content and represent it in a transparent manner
- be deployed in a facilitated group environment
- develop model structure interactively
- incorporate tools to encourage participation and generate commitment to action.

However, they do so in a manner that bears the scrutiny of rigor expected of any science-based approach. Virtually none of the attributes of *soft* OR apply to traditional OR/MS approaches, the latter being increasingly known as *hard* OR (Rosenhead, 2009, S10.) The field of *soft* OR now includes a wide variety of approaches, developed for a range of purposes and applications, some of which could be considered systems approaches. They include:

- Strategic Choice Approach (SCA)
- Strategic Assumption Surfacing and Testing (SAST)
- Soft Systems Methodology (SSM)
- Critical Systems Heuristics (CSH)
- Cognitive Mapping (CM)
- Strategic Options Development and Analysis (SODA)
- Robustness Analysis
- Interactive Planning
- Soft Game Theory including Hyper and Metagames; Drama Theory

The development of such *soft* approaches to address the design limitations of *hard* OR/MS approaches and methods

paralleled the angst of early OR/MS pioneers such as Churchman (1967) and Ackoff (1977; 1979) about the sterility and inappropriateness of overly mathematical approaches for tackling complex social and business problems. Friend and Jessop (1969) developed SCA in the 1960s, as did Mason and Mitroff (1981) with SAST. Checkland and Scholes' (1990) major developments of SSM took place in the 1970s, as did major developments of Soft Game Theory by Howard (1971) and Bennett (1977), and Robustness Analysis by Rosenhead et al. (1972); CM was developed by Eden et al. (1983) in the 1980s; and the major drive to explore hard and soft methods in multi-methodology began to flourish in the early 1990s. However, the work of Munro and Mingers (2002) some 10 years later showed that up to that point, almost all claimed examples of multi-method intervention comprised either all hard or all soft methods, not both in combination.

Such *soft* approaches may provide the opportunity, in Ackoff's terms (1978), to dissolve the problem altogether, to resolve the problem satisfactorily, rather than just optimize or solve a technical problem that is an incomplete or inappropriate representation of the wider, relevant system domain.

Despite the growing evidence that *soft* OR is able to reach problems that traditional or *hard* OR cannot handle, such as organizational and individual behaviors and inconsistencies, *soft* OR is still not well accepted universally. While it is appreciated in the UK and elsewhere, it receives scant coverage in the United States. The hostile reception from journals like *Operations Research* and *Management Science*, which refuse to accept any papers that are "not based on rigorous mathematical models" (Simchi-Levi, 2009, 21), is the topic of current debate (Mingers, 2009a; 2009b).

Much of the *soft* OR story applies equally well to TOC. Indeed, most users of TOC TP would acknowledge their real benefit is in probing the very notion of what the problem is,

why it exists, and what might be the outcomes if it did not exist, before diving into mathematical detail. However, there is also skepticism in traditional circles about whether TOC is a *bona fide* methodology, worthy of publication (Ronen, 2005).

Given the similarity in standing and appreciation offered to nontraditional approaches, one might argue that the time is right for TOC academics and practitioners to unite with academics from across a number of related disciplines, including *soft* OR, to persuade more editors that applications and theoretical developments that do not necessitate a mathematical approach are nevertheless worthy of publication and dissemination. In terms of rigor, TOC does have one clear advantage—the CLR governing the use of the TP provide strict logic protocols that lend rigor to the endeavors of TOC analysts.

Soft OR Methods—Theoretical Underpinnings

In this section, we draw together and reinterpret the prior discussion of *soft* OR methods in the context of the M-B framework, described earlier. We note, in particular, that soft approaches have been designed and developed to assist with all phases of problem intervention—*appreciation, analysis, assessment, action*—but especially as they relate to matters in the *social* and *personal* domains of the M-B classificatory system.

[Table 23-4](#) provides examples illustrating how two *soft OR* methods, namely SSM and CM (Cognitive Mapping), map to the M-B framework. The “+” symbols indicate the relative extent to which each tool is purposively designed to attend to each phase of problem intervention in each of the problem dimensions. We note, for example, that SSM and CM have not been expressly designed to contribute to the

analysis and understanding of underlying causal relationships in the material world—although their use may well contribute to doing so.

As may be inferred from commentary on the characteristics of *hard* OR/MS approaches, *soft* approaches have been developed to assist with situations where problems are not well-defined and where goals, local or global, are not necessarily understood, known, or accepted; where multiple stakeholders are involved; where desired outcomes cannot be guaranteed by action; and where success is ambiguous and the definition of success may need to be negotiated. As a consequence, *soft* approaches meet a need to facilitate learning about a problem, its constituency and constituents, their customs, practices, and ways of thinking; that is, they also seek to explore and accommodate a range of views, worldviews, values, and objectives without reducing them to a single measure, and they seek to encourage the active involvement, engagement, and commitment of stakeholders (Mingers, 2009b).

These latter features of the soft approaches are reflected by multiple “+” signs emphasizing the level of their contribution to the personal and social domains of problem context, across the whole spectrum of phases of problem intervention in the M-B classificatory framework. By contrast, hard approaches tend to be situated to provide *analysis* and *assessment* within the *material* world. Furthermore, the notion that *soft* approaches meet a need to facilitate learning about a problem aligns with Checkland and Scholes’ view that *soft* systems approaches can or should also be conceptualized as “learning systems” (1990, A8).

Phase Dimension	Sensibility: Awareness, Empathy, Appreciation of . . .	Analysis Understanding and Synthesis of . . .	Appraisal, Evaluation, Assessment of . . .	Purposeful Action Choices to . . .
Personal	Individuals' ideas, beliefs, meanings, emotions, aims, needs and wants SSM ++ CM ++++++	Different perspectives, perceptions and worldviews— Weltanschauung SSM ++++++ CM ++++++	Alternative conceptualizations and constructions of reality SSM ++++++ CM +++	Create common ground, and consensus about ideas, states, etc. SSM +++++ CM +++
Social	Social context, norms, practices, relationships, power relations SSM ++ CM +++	Misperceptions, misrepresentations, distortions, conflicts of interests SSM CM +++	Alternative means of critiquing, contesting or modifying power relationships SSM CM +++	Generate understanding and empowerment to effect desired relationships, states, etc. SSM +++++ CM +++
Material	Physical context and relationships SSM ++ CM	Underlying causal relationships and structure SSM CM	Alternative physical and structural arrangements SSM CM	Identify, select and implement best alternatives SSM CM
	←Problem-Solving→		←Decision-Making→	

Source: Adapted from Mingers (2000).
Simon et al. (1987) facilitate the comparison of the M-B 4 phases with Simon et al.'s conceptualization of problem-solving and decision-making.
CM is Cognitive Mapping.

TABLE 23-4 Mapping of SSM and CM

For instance, and for [Table 23-4](#), we note that within the *action* and *implementation* phase that both SSM and CM are designed to seek out and seek accommodation of disparate views/or consensus, and to seek enlightenment and empowerment for problem constituents as well as the problem owner and analysts. In stark contrast to *hard* OR/MS methods, we further note that the primary purpose of such *soft* methods is to understand better, not necessarily to identify best alternatives.

CM, similarly to SSM, seeks to effect a representation of how individuals view a problem, what it means to them, and how they make sense of it. The axiology or purpose of CM is to surface and understand these beliefs in order to generate consensus about possible strategic action. In contrast, we reaffirm that hard OR/MS relates mostly to the material world and focuses on *analysis* and *assessment*, leading to action in that domain. However, both *hard* and *soft* methods and methodologies may also be described as systems approaches. The nature of systems approaches and methodologies will be explored further in the next section.

Systems Approaches

Systems approaches to problem-solving typically conceptualize “problems” as existing within a notional whole or synthetic system, where a system can be defined as any grouping of people, events, activities, things, or ideas, connected by some common reason or purpose (Senge, 1990). As such, many systems can be best described as notional. In general, we can describe systems as being natural, for example, ecological systems; as being designed, for example, a car or an organization; or as being a human activity system, for example, a sports team or an ad hoc work group. Systems thinking attempts to reflect and illustrate the importance of holism, of boundaries, of feedback, of reciprocal relationships, and the notion that, say, activities or events, while perhaps separated by distance and time, cannot be understood in isolation, but instead need to be understood in terms of the patterns of relationships that create them and the patterns of behaviors that emerge from those relationships.

Systems thinking entails, above all, a sensibility about matters systemic (Espejo, 2006); that is, consideration of the big picture, the need to think holistically, to consider the

whole as a network of relationships of interconnected parts or subsystems, and the need to understand feedback. Systems thinking, according to Senge (1990), involves learning to recognize structures that occur repeatedly—a notion that accords with TOC practice.

We often seek to understand a problem or problem situation by taking a Cartesian reductionist approach to analyze and understand its identifiable “component parts.” However, systems thinking reflects a recognition that to understand a problematic situation fully, or why a problem exists, and persists, requires the problem to be situated within a wider context, a notional whole or system, and then to understand how the parts of that system relate or contribute to the whole—which, in itself, is an act of synthesis or synthetic systems thinking. Indeed, such conceptualization of problem situations as systems is an act of systems thinking.

In general, we may make a useful distinction between representing reality as systems, using systems language and protocols, and inquiring of what we regard as reality using systems approaches (Senge, 1990; Checkland and Scholes, 1990). The latter notion is that by examining situations using systems frameworks as learning frameworks, and by using the systems concepts of holism, boundary, feedback, etc., one can gain an understanding of complex situations where seemingly insignificant events can catalyze the playing out of complex relationships that generate unpredictable, unanticipated emergent behaviors and outcomes, which cannot be attributed to any single causal event.

Senge (1990) regards systems thinking as a discipline for seeing wholes; as a framework for seeing inter-relationships rather than events, and for seeing patterns; as a set of principles; and as a sensibility for “the interconnectedness that gives living systems their unique character” (69). In

adopting a systems approach, one is therefore less likely to be reactive or over-reactive to current or local events or outcomes, where such over-reaction may potentially exacerbate undesired problems elsewhere. As a corollary, we may become more sensitized to patterns of change, to the impact of change, and of the systemic influences whereby even a positive change in one area of a system may lead to adverse effects elsewhere in another part of the system. As such, systemic sensibility to such possibilities may likely reduce the tendency to act and think suboptimally and, as Senge (1990) suggests, to be generative in terms of creating systemic structure that leads to sustainable and desirable outcomes.

While Jackson (2000) has asserted that systems thinking is, indeed, a new paradigm which could revolutionize management practice in the 21st century, Senge (1990) sees systems thinking as a discipline to change patterns of thinking. In recognizing the evolving broad church of systems approaches, Jackson (2000) has commented on the need to recognize communality and complementarity in the methodology and purpose of such approaches. He offers “critical systems thinking” as a coherent framework to unite diverse systems approaches, including chaos and complexity theory, the learning organization, SD, living systems theory, SSM, interactive management, interactive planning, total systems intervention, autopoiesis, management cybernetics, the viable system model, operations research (hard and soft), systems analysis, systems engineering, general system theory, socio-technical systems thinking, the fifth discipline, social systems design, team syntegrity, and post-modern systems thinking. However, Jackson, like most systems thinkers, fails to mention TOC as belonging to this broad church, despite most TOC authors (Goldratt, Dettmer, Scheinkopf, Cox et al.,

2003) labeling TOC as a systems approach and stressing the importance of taking a systems view.

As such, we can view, for example, Checkland and Scholes' (1990) SSM or Beer's (1985) *Viable Systems Model* (VSM) of organizational structure and design as enquiry systems, as learning systems, where the methodology or model provide the conceptual framework to guide our inquiry and learning about the situation or organization at hand. In both these cases, the notion of purposeful systems looms large in the mode of inquiry. In SSM, assumptions about the nature and purpose of the system being examined are captured in a "root definition" stated in terms of its **C**ustomers, **A**ctors, **T**ransformation, **W**eltanschauung, **O**wners, and **E**nvironment (**CATWOE**). Any attempt at problem solving, therefore, takes place in the context of the system definition and purpose. Similarly, any use of the VSM to explore organizational effectiveness or the effectiveness of organizational design is made in the context of defined organizational purpose.

We note the similar importance of system goals or purpose in the effective use of the 5FS approach, the EC process, and more generally within TOC of identifying *What to Change to?* We also note the underpinning assumption in the development of GEC that the seemingly initial different worldviews of analysts can be accommodated in a single generic cloud. These matters beg the question, notwithstanding the importance of defining the system goal, of whether the system goal can be objectively defined or whether the system goal remains an ill-defined phenomenon whose definition and description varies according to the questioner/observer.

Other Decision-Making Tools

There are many other cause-effect based tools and decision-making models available, each of which has its advantages. For example, the *Theory for Inventive Problem Solving* (Altshuller, 1973), known by its Russian acronym, TRIZ, provides a useful method for generating solutions to dilemmas (or contradictions), and may be usefully employed in tandem with the EC (Mann and Stratton, 2000; Dettmer, [Chapter 19](#) this volume). *Why-Because Analysis* (WBA) provides an alternative method of constructing a fault analysis to root cause, and has also been used in tandem with and compared with TOC⁵ (Doggett, 2004; 2005; Zotov et al., 2004). In addition, process mapping has also been used to aid CRT construction, and other techniques, including Lean, quality management, and process engineering, have been found to be mutually supportive of, and certainly not mutually exclusive of TP tool use (Watson et al., 2007).

Lessons for TOC from the Literature

Issues Emerging from the TOC Literature

Ronen (2005), in his guest editorial introducing the special issue on TOC published in the *Journal of Human Systems Management*, bemoans TOC's low profile in academic research journals, and offers some reasons why this may be the case:

- TOC is heuristic oriented, in line with Simon's et al. (1987) "satisficing." Many academic journals prefer process-optimizing, quantitative approaches, while the goal of TOC is simplicity.

- TOC processes are cause-effect driven. Academic journals prefer field studies or empirical data.
- TOC originated in practice—not enough academics have been exposed to its full contribution.
- TOC is often misperceived as a simplistic toolkit that does not need thorough research.
- TOC is viewed as a cult and thus inaccessible to the academic community.

Ronen called on academics to apply academic methodologies to TOC concepts and confirm or improve its methods, and to apply academic rigor to research on TOC. More specifically, Watson et al. (2007), in their “Silver Anniversary” review of TOC, identified two common problems with TP:

- The reliance on subjective interpretation of perceived reality and the qualitative nature of the subject matter makes the tools inherently unreliable, which leads to a perceived lack of reliability and validity in TP analyses.
- The TP tools are criticized for not being user friendly.

These matters are outlined further in the next section.

The Nature of the TOC Literature Vis-à-Vis Other Literatures

We now comment more broadly on the nature of the TOC TP literature, its distinctiveness, and similarities and differences with the literature of other OR/MS and other problem-solving methodologies. In order for TOC to increase its visibility—and acceptability, especially within academia—there is a suggested need for TOC practitioners and academics to target publication of refereed journal articles

and book chapters in edited volumes, to counterbalance the many non-refereed books and conference papers on TOC, to build that greater visibility and credibility.

However, and second, there is no one journal ideally suited for TOC or the TOC TP, although there are several obvious journals for production applications, projects, etc. As a result, published work as articles is spread across a range of journals, promoting widespread coverage but possibly reducing the impact that may come from there being a concentration, or home, in particular journals. Furthermore, as Kim et al. (2008) noted, the journals that published articles on TOC TP were generally those with lower impact factors, so targeting higher impact journals would be desirable.

Third, and unfortunately, while TOC would appear to share many characteristics with OR/MS, other systems methodologies, and *soft* OR, proponents of these methods do not generally consider TOC to be one of their kind, due to a lack of awareness, understanding, or through a more deliberate choice to exclude—an irony indeed, when one considers the history and plight of *soft* OR in seeking to gain acceptance in the more traditional OR/MS and journal mainstreams.

If the TP are to gain more recognition within academia, we must break into or join the journal mainstream, and despite the previous comments, the lowest barriers to entry may be through links with these comparator or peer disciplines. Indeed, Ronen (2005) has suggested that many TOC practices have their roots in well-accepted and well-established OR/MS concepts, which facilitates their multi-methodological use, for example, combining the 5FS and the mathematical programming approach in multi-methodological use. There have been several papers comparing TOC with LP, mostly showing congruence, but also showing advantages of using TOC. Indeed, there can be significant synergies in combining TOC with OR methods, as

argued by Mabin and Gibson (1998), echoing criticism by Zeleny (1981) and Gass (1989) of naive usage of LP in relation to the management of constraints.

The reviews of the TOC and TOC TP literature reported here have been complemented by a useful series of retrospective/state-of-the-art reviews conducted for the 50th anniversary conference of the Operational Research Society, York, 2008, many of which are published in Brailsford et al. (2009). These included reviews on *soft* OR and PSMs (Rosenhead, 2009), Systems Thinking (Jackson, 2009), and healthcare (Royston, 2009).

What is noticeable and notable in these reviews is that TOC is absent, unrecognized, or excluded from descriptions of OR, *soft* OR, PSM, and systems methods as presented in such reviews. For example:

The notable successful work at Radcliffe and Horton Hospitals by the Goldratt Group has been reported in Umble and Umble (2006) as well as in *The Oxford Story*, by Dr. Eli Goldratt, available on various websites,⁶ but this work is not included in the “One hundred year review of OR in health,” despite it being published in a prominent OR journal.

TOC is not mentioned in the *soft* OR discussions of Mingers (2009a; b), despite being linked and compared in the *Omega* publication of Davies et al. (2005), and despite sharing many seeming commonalities in domains of application.

TOC is not generally referred to in OR texts, except sometimes for a small section on OPT, constraint management, or synchronous manufacturing aspects. Even in Operations Management textbooks, there is usually just one chapter on TOC, with a few notable

exceptions such as Cox et al. (2003), which approaches the subject using TOC as the overarching framework.

Even though TOC work has been published in OR/MS and systems journals, it appears that TOC has yet to be considered mainstream or to provide a mainstream contribution to any of these disciplines. Nevertheless, the TOC community can do more than await recognition. However, before taking action appropriate to gaining that recognition, there would be benefit in conducting a self-audit. The following section offers some suggestions about matters that deserve consideration.

Suggested Topics for a Self-Audit of TOC

In the following subsections, we suggest much can be gained from a self-analysis of TOC as a field of learning and as a profession, and the role of TOC academics within the profession. The self-analysis should encompass the strategic role of publication outlets in gaining recognition for TOC, and which may best serve academics, practitioners and the wider community provided for by TOC. In a subsequent section, we also suggest that much can be gained by understanding the nature of TOC as a methodology.

TOC as a Profession

The TOC community is not alone in its experiences. There is a sense that much can be gained from looking inward *and* outward, and it can learn from the concerns and experiences of other professional groups. For example, OR in the United States has suffered from what Abbott (1988) termed “professional regression”—a process by which professions withdraw into themselves (Rosenhead, 2009, S13, quoting Corbett and van Wassenhove, 1993). Furthermore, status rankings, internal to the profession,

based on the knowledge system that gives a profession its special claim, tend to be correlated with remoteness from practical concerns and implementation. Rosenhead asserts that the 1988 CONDOR report illustrated this tendency in OR.

At present, the TOC profession seems safe on this latter tendency, as TOC developments are strongly practice-based or practice-oriented (Inman et al., 2009). We suggest that the TOC profession should be aware of the risk of professional regression, but acknowledge that there is an inherent dilemma. On the one hand, if TOC wishes to gain credibility and recognition within and from other peer disciplines, it needs to conform to the academic rigor and norms of those peer disciplines. However, in order to do so, the TOC community must submit its body of knowledge to scrutiny using the same academic norms and protocols to which other academic peer groups are subject. If TOC fails to build support from other peer disciplines, it can run the risk of “professional regression.” However, if TOC seeks to gain such support by uncritically adopting the methods of other disciplines, it may jeopardize TOC’s focus on practical aspects that traditionally motivate TOC proponents and that fuel most of the developments within the TOC community.

Identity and the Strategic Role of Publication Outlet

The previous section has made implicit reference to the important issue of identity—both self-identity and the identity projected to others. In considering these matters further, we need to consider more broadly why TOC has not become accepted in the mainstream, and more specifically, why TOC is rarely mentioned in the academic and journal mainstream. One may suppose that TOC is not recognized as being OR or *soft* OR because of its very distinct parentage, and that many may still think of TOC as a

scheduling or manufacturing method. We suggest that much may be gained from spreading the message appropriately, demonstrating that TOC is more than a set of tools for operations management. Constructive illustration of the TP in complementary or multi-methodology work, in other domains of application, may help build awareness and acceptance of the TP.

Even though TOC is not considered to be hard or *soft* OR, or a systems method by proponents of those disciplines, TOC TP contributions are already finding favor in the UK and in the European more practically oriented OR and systems journals such as the *Journal of the Operational Research Society* (JORS) or the international OR federation's *International Transactions in Operational Research* (ITOR). The journals *Human Systems Management* (HSM) and *International Journal of Production Research* (IJPR) have both published special issues on TOC. Maybe the time is right to explore similar outlets such as the *Journal of Operations Management* (JOM) following the success of Watson et al. (2007), *European Journal of Operational Research* (EJOR), *Interfaces*, and other *INFORMS* journals, especially given the currency of debate about *soft* OR; and especially so, if one may argue that TOC could be considered either as constituted in the same manner as other *soft* OR methodologies or considered as part of the *soft* OR community or domain. Indeed, since more support has been shown overtly, at the time of writing, for *soft* OR in the U.S. OR community, there may also be an increased readiness for publishing TOC papers within the U.S.-based journals.

Role for TOC Academics/Researchers

We may infer from the broader discussion of the literature, and prior comments, that there is a need for TOC academics and researchers to remain connected with practice, while

building academic credibility through rigorous research. In this role, we may suggest that TOC academics should aim:

- to link, interpret, and comment on TOC knowledge and practice from an objective perspective;
- to further develop TOC knowledge in ways that embed TOC into extant academic disciplinary knowledge and leverage off those other disciplines;
- to enhance the academic qualities of TOC knowledge, and the status of TOC in academia; and
- to begin a dialog with TOC practitioners on these matters, in the hope that they will find such dialog valuable and useful, as they continually reflect on their practice as part of their own continuous improvement processes.

The next section provides a first step in satisfying these aims, in as much as it seeks to summarize, reinterpret, and build our understanding of the nature of the TOC TP, and TOC as a methodology.

The Nature and Use of the TOC Thinking Processes Revisited

Here, we subject the TOC TP and TOC as a methodology to examination using the classificatory frameworks of Mingers and Brocklesby (1997) and Mingers (2003). As such, we heed Ronen's (2005) call for more rigor in the TOC domain, and his call to establish the credibility of TOC, by providing an external perspective using the frameworks in a transparent and rigorous fashion. As a consequence, we also work toward Ronen's goal to close the gap between TOC and the academic world.

In the following section, we first draw on the M-B framework (1997) to provide an alternative perspective of different TOC methods and TP tools, by clarifying their role, function, and purpose. We are then able to relate the methods and tools, and broader TOC methodology to problem content and problem-solving activity—in order to provide a basis for selective comparison with traditional methods. In the subsequent section, we seek to surface and clarify the underpinning philosophical assumptions that support TOC TP, other TOC methods, and TOC as a methodology.

Understanding the Relationship of the TOC TP to Problem-Solving Activity

In [Table 23-5](#), following the M-B classificatory approach, we characterize a selection of the TOC TP tools used within TOC using the descriptions of each of the tools and methods, as a basis for such characterization and classification (see full set in Davies et al., 2005).

The bolding of the TP tool name reflects our view that the tool was developed and designed for purposeful use in a particular phase of the problem-solving process, while the number of “+” signs indicates the extent to which the tool was designed to meet such purposes.

In an illustrative interpretation of the characterization, we note, for example, that the mapping of EC activity to the modified M-B framework (see [Table 23-5](#)) demonstrates how the EC method can provide an effective bridge from the problematic current situation to the desired future by contributing to all phases of intervention, but not necessarily across all problem domains.

Similarly, we note that the set of tools and methods of TOC are designed in a way that they may contribute across

all phases of problem-solving activity including what we refer to as *action* or implementation.

In addition, the tools directly target or deliver on all but one of the cells in the M-B grid (see Tables 23-5 and 23-6), namely, the *appraisal and evaluation of* means of critiquing, contesting, or modifying power relationships in the social domain. In particular, we indicate, by the increasing darkness of shading in [Table 23-6](#), the relative extent to which the collective set of TP tools is purposively designed to attend to each phase of problem intervention in each of the problem dimensions.

	Sensibility: Awareness, Empathy, Appreciation of . . .	Analysis Understanding and Synthesis of . . .	Appraisal, Evaluation, Assessment of . . .	Purposeful Action Choices to . . .
Personal	Individuals' ideas, beliefs, meanings, emotions, aims, needs, and wants CRT EC ++++ FRT PRT +++++ TRT ++++ TOC as a meta-methodology ++++	Different perspectives, perceptions and worldviews— Weltanschauung CRT EC +++++ FRT PRT TRT TOC as a meta-methodology ++++	Alternative conceptualizations and constructions of reality. CRT ++ EC +++++ FRT PRT TRT TOC as a meta-methodology ++++	Create common ground, and consensus about ideas, states, etc. CRT EC +++++ FRT PRT +++++ TRT TOC as a meta-methodology +++++
Social	Social context, norms, practices, relationships, power relations CRT ++	Misperceptions, misrepresentations, distortions, conflicts of interests CRT	Alternative means of critiquing, contesting, or modifying power relationships	Build understanding and effect empowerment to create desired relationships,

	EC FRT PRT +++++ TRT +++++ TOC as a meta-methodology ++++	EC +++++ FRT PRT TRT TOC as a meta-methodology ++	CRT EC FRT PRT TRT TOC as a meta-methodology nil	states, etc. CRT EC +++++ FRT PRT +++++ TRT TOC as a meta-methodology +++++
Material	Physical context and relationships CRT +++++ EC ++ FRT ++ PRT TRT TOC as a meta-methodology ++++	Underlying causal relationships and structure CRT +++++ EC +++++ FRT +++++ PRT TRT +++++ TOC as a meta-methodology +++++	Alternative physical and structural relationships CRT EC FRT +++++ PRT +++++ TRT +++++ TOC as a meta-methodology ++++	Identify, select, and implement best alternatives CRT ++ EC +++++ FRT +++++ PRT TRT +++++ TOC as a meta-methodology +++++

TABLE 23-5 Mapping Methodologies—TOC TP

	Sensibility: Awareness, Empathy, Appreciation of . . .	Analysis Understanding and Synthesis of . . .	Appraisal, Evaluation, Assessment of . . .	Purposeful Action Choices to . . .
Personal	Individuals' ideas, beliefs, meanings, emotions, aims, needs, and wants	Different perspectives, perceptions and worldviews— <i>Weltanschauung</i>	Alternative conceptualizations and constructions of reality	Create common ground, accommodation, and consensus about ideas, states, etc.
Social	Social context, norms, practices, relationships, power relationships	Misperceptions, misrepresentations, distortions, conflicts of interests	Alternative means of critiquing, contesting, or modifying power relationships and structures	Build understanding and effect empowerment to create desired relationships, states, etc.
Material	Physical context and relationships	Underlying causal relationships and structure	Alternative physical and structural relationships	Identification, selection, and implementation of best alternatives

Adapted from Mingers and Brocklesby (1997) and Davies, Mabin, and Balderstone (2005).

TABLE 23-6 Alternative Mapping of TOC as Meta-Methodology

In explanation of such categorization, we refer to the protocols and criteria of the M-B classificatory system relating to purposeful design. We note, for example, that while some *soft OR* methods were expressly designed and developed with the purpose of setting out to contest or change power relationships and structures, we cannot say that the TOC TP were designed for that specific purpose. That fact notwithstanding, TP tools have and may be used to attend to such matters successfully.

Indeed, the TOC TP may not address such issues unless diagnosis (using, say, the CRT) points to the power structure as being a core problem, or unless the power structure is seen to be an obstacle during the development of the PRT. Even though such challenge to power structures may be an emergent property of the TOC approach, since TOC does not aim to do this from the outset, nor is it a natural common outcome, we have left this box unshaded to maintain consistency with Mingers' classificatory approach—which is that the classification of an activity requires that it is deliberately designed for that phase of intervention. Nevertheless, we may conclude that the characterization demonstrates that the TOC TP comprise what Dettmer calls the “complete package” and what we call a methodological set or meta-methodology.

The next section demonstrates how the related classificatory framework of Mingers (2003) may give rise to complementary insights.

The Philosophical Basis of the TOC TP

In [Table 23-7](#), we provide an alternative characterization of each of the TOC TP tools and the 5FS method. In doing so,

we again draw on the brief descriptions of each of the tools and methods as a basis for the characterization of the underpinning philosophical assumptions, using the classificatory system of Mingers (2003).

	Functionality	Axiology		Ontology	Epistemology		
TOC Technique or Tool	... that has a function to for specific users including for the purpose of having made assumptions of what exists ...	and representing and modelling what exists, via ...	using available information such as ...	obtaining such information by ...
Five Focusing Steps	... Identify and manage constraints on continuous improvement	Participants, decision-makers and implementers, stakeholders.	... improving global performance long-term	Constrained performance, barriers to improved performance	A process of identifying and examining constraints on performance	Objective facts, opinions, logic relations, judgements, desired outcomes, necessary actions	Observation and measurement of real world, judgement and opinion
Current Reality Trees	... search for root causes, and explain how these lead to problem symptoms	Decision-maker, analyst, consultant, facilitator, participant	... discovering root causes to problems	Problems, symptoms, cause-effect relations	The mapping of cause-effect/logic relationships	Objective facts, subjective opinions, logic relations, perceptions, judgements, patterns of behavior	Observation and measurement of real world, logic relations, judgement and opinion
Evaporating Clouds	... represent explicitly one or more persons' conflicting views	Analyst, participant	... surfacing and understanding individual beliefs, synthesising competing viewpoints, explaining how these lead to conflict, generate actions	Individual beliefs about competing views and the assumptions underlying these views of different stakeholders	The mapping of seemingly diametrically opposed viewpoints, objectives, necessary conditions, underlying assumptions, and relevant stakeholders	Options, stakeholder viewpoints, and their interests	Interviews, discussion, argument, debate with participants, analyst's reasoning
Negative Branches (NBRs)	... Identify possible side effects and actions to prevent them	Participants, decision-makers and implementers, stakeholders	... identifying causal actions required to prevent undesirable side effects	Existence and elimination of undesirable side effects of proposed action	The mapping of cause-effect/logic relations and side effects from actions	Objective facts, subjective opinions, logic relations, judgements, side-effects and actions to overcome them	Observation and measurement of real world, judgement and opinion

Future Reality Trees	... determine effects and outcomes following from proposed actions and solutions	Decision-makers, analyst, consultant, facilitator, participant	... showing how actions lead to desired outcomes	Problems, actions, desired outcomes, outcomes, cause-effect relations	The mapping of cause-effect/logic relationships	Objective facts, subjective opinions, logic relations, judgements	Observation and measurement of real world, judgement
Prerequisite Trees	... surface and list obstacles and necessary corrective actions to achieving desired outcomes	Participants, decision makers and implementers, stakeholders	... mapping the necessary sequence of actions required to achieve desired outcomes or target	Existence of implicit obstacles to achieving desired outcomes	The mapping of necessity relations between necessary actions to overcome obstacles in the form of a map	Obstacles, and actions to overcome them, logic relations	Viewpoints, intuition, judgement
Transition Trees	... Identify required actions to generate desired outcomes and results	Participants, decision maker, implementers, stakeholders	... creating an action plan to achieve desired outcomes	Problems, actions, desired outcomes, outcomes, cause-effect relations	The mapping of cause-effect/logic relations in the form of a map, actions, desired outcomes	Objective facts, subjective opinions, logic relations, judgements, desired outcomes, actions to achieve them	Observation and measurement of real world, judgement and opinion

Adapted from Davies, Mabin, and Balderstone (2005).

TABLE 23-7 Framework for Characterizing the Philosophical Assumptions underlying TOC Methods

We note that when the underlying assumptions and purpose are to be presented in this manner, we need to gain clarity about for what purpose and how the tools may be best used, and we may then develop realistic expectations

about the use of the tools. In addition, we also note and foreshadow the scope for complementary use of the tools with respect to addressing multipurpose or multiobjective problem situations.

In the following section, we will re-examine the tools and their purposes in terms of their contributions to the different phases of intervention in the problem-solving process. It is worth restating here that even though the tools and methods are often used on their own for day-to-day problems, they are often used in combination for more infrequent and complex situations as well (Kim et al., 2008). The nature of such use, and reasons for its success or failure, can be explored appropriately by reference to the characterization of TOC tools presented in [Table 23-7](#).

[Table 23-7](#) captures and represents succinctly the defining nature and nuanced purposes of the TOC TP. In doing so, it makes explicit, somewhat ironically, the often unstated, sometimes unrecognized philosophical assumptions that underpin the TP, the associated TP tools, and their use. Some such assumptions relate to beliefs about what exists—cause-effect relations—and what could be—continuous or breakthrough improvement—and are *ontological* in nature. Other assumptions relate to the nature of information available, how we may access such information, and how we represent and process it, via causal logic trees. These are *epistemological* in nature. Similarly, other assumptions or beliefs relate to what we may expect a TOC tool “to do,” and to its *axiological* nature; that is, for whom the analysis is being conducted, and for what purpose the tool will be used. As such, [Table 23-7](#) provides a different perspective on the TP tools, and on their development and use, especially the need to be cognizant of and in tune with the philosophical assumptions when seeking to use the tools appropriately and effectively.

Summary Insights from Classificatory Mapping of the TOC TP

Recognizing and Understanding TOC as a Systems Meta-Methodology

The mapping of the various TOC TP tools and methods to the Mingers and M-B frameworks shows that they not only overlap or substitute for each other to some degree, in terms of purpose and underlying philosophical assumptions, but that they may also be complementary in nature. Indeed, whereas we may expect similar insights to arise from more than one method or frame, in general, there may also be new insights about the problem, and how it should be tackled, arising from each.

As a result, we suggest there will be, in most cases, no one best model, method, or methodology; and as such any implicit search for a “best-fit” model or method should be surfaced explicitly and abandoned. If so, the pragmatic adoption of what may then be a multi-method or multi-methodological approach accords with Burrell and Morgan (1979) and Brocklesby (1993) in their discussion and acceptance of the efficacy of multi-paradigm and multi-methodology development.

Seldom are any of the TOC methods and tools used in isolation. Certainly, for complex problems, several tools may be, and are often used as problem-solving intervention moves through the stages from diagnosis to implementation (Kim et al., 2008). Using the conceptualizations of the M-B framework, we recognize that TOC methods are often used as complements to broaden or heighten, for example, the *appreciation* phase of intervention, or to complement *analysis* and *assessment/evaluation* with a stronger *action/implementation* phase.

When the full set of TOC tools and methods considered here are mapped to the M-B framework (see Tables 23-5 and 23-6), we note how these methods may comprise a multi-method approach, attending to almost all phases of

intervention across all dimensions of the problem domain. Consequently, they can be regarded as a methodological set. We also note the potential for further discussion of whether the broad umbrella of TOC can be regarded as a meta-methodology, a meta-framework, or a multi-methodological approach. We also note the irony in the juxtaposition of the benefits of such potential discussion, and the lack of deep understanding about TOC that prevails.

Observations—A Lack of Deep Understanding Prevails about TOC

There has been an unfortunate lacuna in the TOC literature relating to the nature of methodology and of methodological developments. Consequently, there has been an absence of the necessary base for critical reflection about methodology-in-use. Invoking Argyris and Schön's (1974) notion of double loop learning which stresses the importance of reflection about experiences for learning to take place (Schön, 1983; Kolb, 1984), we would argue that TOC practitioners are no different from others in needing to be critically reflective about their experiences of using TOC. Such critical reflection is a necessary condition for a deeper understanding of TOC by its users.

One example of a lack of reflection, or of a shortcoming in TOC methodology, relates to the systemic nature of the TP. In particular, it relates to the minimal presence or relative absence of a critical component of systems thinking; that is, feedback and feedback loops. Whereas systems representations embodying CLDs actively search for feedback loops in the depiction of cause-effect relationships, the process by which, for example, a CRT is developed, linking problem symptoms in a chain of causal relations to the root cause, to some extent, mitigates against the identification of feedback loops.

In the building of a CRT, feedback loops tend to be added in the latter steps of the process, almost at the conclusion of building the tree. Moreover, such loops are typically labeled as “negative feedback loops” because they refer to the continuing and “negative” unwanted or desirable nature of the situation being described.

However, while TOC’s definition and usage of feedback loops is unambiguous in the context of TOC, and used with consistency within the TOC community, such definition and usage are unnecessarily out of step with the rest of the systems community. We suggest that a change in process and definition should be contemplated by the TOC community. In other systems methodologies, any feedback loop that reinforces an effect is termed a “positive feedback loop.” Indeed, in most systems methodologies, both “vicious” and “virtuous cycles” are conceptualized and labeled as positive feedback loops—the simplest example being that of two variables that act on each other in a mutually reinforcing manner—with each variable causally effecting and being causally effected by the other to create greater effects of a positive or negative nature. By contrast, a negative loop is one that moderates an effect, that brings a variable back on course in the sense that it incorporates one or more cause-effect relationships that bring a system back to a desired state, like a thermostatically controlled air conditioning system. Relabeling and redefining the TOC loops accordingly would facilitate a shared understanding and acceptance of TP diagrams, and the TOC approach more generally, by other systems communities.

Exploring the link between TOC and other systems methodologies can also enhance understanding of problem situations. Indeed, we have argued elsewhere (e.g., Davies and Mabin, 2009) that each of the EC and CLD representations can be enhanced by multi-methodological use to display relationships, not only using *necessity* logic, but also—with appropriate explanatory intermediate

variables emanating from assumptions underlying the EC—using *if-then* sufficiency logic. Such examples illustrate beneficial developments in TOC methods that may be sought over time to improve TOC as a methodology, and to enhance the use of particular TP tools.

Ronen's comments (2005) also suggest that, regardless of such shortcomings, it is necessary to establish the credibility of TOC as a methodology within academia. In his early writings, Goldratt (1990b, 23) described the development of scientific theories as a progression through classification, correlation, and causation stages. Here we have provided the classificatory frameworks that form a basis for understanding how TOC TP methods and methodology within TOC are constituted.

Summary

What Has Been Covered in This Chapter

This chapter has provided an overview of the TP that has addressed their conceptual, philosophical, and methodological foundations, alongside discussion of the TP use and practice. As such, we have been able to reflect on the need for the TP; on the design and purpose of the TP; on their effective use in practice; on reasons for their existence and effectiveness; and we have done so in order to effect a consolidation of our understanding of the TP that may serve as a platform for future developments and use.

In doing so, we have provided a supporting rationale for the existence of the TP by explaining how they meet needs of a methodological and practical nature that are not addressed by other problem-structuring and problem-solving methods, for example, those of OR/MS. We have also

suggested that there is a need to explore how the TP may be used in multi-method and multi-methodological intervention with, say, OR/MS or systems methods, with other TP tools, and with other TOC methods. In addition, we believe that building such links and bridges with cognate fields and disciplines through multi-methodological intervention, exploiting identified synergies, may well serve to gain further acceptance for TOC within those cognate fields, through the building of communities of practice, with, say, those embracing systems and *soft* OR methodologies.

Furthermore, we respond to the call for the domain of TOC practice, and for TOC as an academic field of inquiry, to gain further recognition from cognate professional groups and academia, by suggesting further engagement in research on TOC methods and practice that satisfy the demands of professional and academic rigor, open doors to highly regarded publication outlets, and to acceptance of TOC as a *bona fide* academic endeavor. The following section addresses these matters.

Findings and Recommendations

There is seemingly ample evidence of how diverse issues and problems can be tackled effectively using a variety of Goldratt's TOC tools, principles, and methods—from the simplistic product mix algorithm, the 5FS, *Drum-Buffer-Rope* (DBR), *Buffer Management* (BM), *Critical Chain* (CC), the EC, to the suite of TP (Rahman, 1998; Kim et al., 2008; Mabin and Balderstone, 2000; 2003; Mabin and Davies, 2003; Inman et al., 2009; Watson et al., 2007).

The review of Kim et al. (2008), reported here, revealed specific publication and research gaps, and some common future research topics and approaches have emerged. First, no work has been published that relates to critical success factors or necessary conditions underpinning the effective

implementation of the TP. Given the empirical importance of measuring and comparing the rate of success or failure with other business improvement approaches, such as ERP, Lean, or Six Sigma, it is a little surprising that the publication of research on this topic has attracted minimal attention among TP academics. Further investigation of critical success factors and common problems in the application of the TP is definitely required.

Second, in order to provide practitioners and academics with a critical evaluation of TP tools-in-use, the lack of published empirical work on the effectiveness of TP applications must be addressed. Inman et al.'s (2009) cross-sectional analysis, using structural equation modeling, to examine the links between elements of TOC use, TOC outcomes, and organizational performance has provided illustration of analysis not previously attempted, as a means of filling such gaps. Further empirical studies of both cross-sectional and longitudinal nature, across industries and applications, and over time, would be appropriate, in as much as they would promote the testing of research hypotheses and would strengthen the TP knowledge base. In particular, such research could be directed toward identifying and measuring performance before and after TOC implementation.

Third, the literature reveals an ongoing discussion and critique regarding the philosophical underpinnings of TOC as a methodology. One apparent limitation with the use of the TP is that their use appears to be problem driven; they are applied only when there is a "problem" (Tanner and Honeycutt, 1996; Antunes et al., 2004). The review suggests that there is an unmet need for studies exploring how TOC methods can be applied, not just in problem situations, but also in situations that are problematic in a positive rather than a negative sense. This approach reflects a paradigm shift that has been termed "blue ocean" strategy. Kim and Mauborgne (2005) argue that most companies need to

create blue oceans of opportunities. They show how a company can create a blue ocean by changing its strategic thinking and using a systems approach. As such, Kim et al. (2008) recommended that more consideration should be given to how the TP could be applied in situations where positives are renewed and advanced, rather than just responding to negatives, or to a need to eliminate or ameliorate problems. The recent development of S&T and their application in situations where “stretch” strategic goals may be set, either to ameliorate or eliminate negatives or to pursue positives, would appear to address this gap. S&T trees were not found in the peer-reviewed literature, but are discussed in [Chapters 15](#), [18](#), [22](#), [25](#), [31](#), and [34](#) in this volume.

Furthermore, the overview presented here suggests that there is scope for building on the considerable work that uses the non-TP part of TOC, to test the impact of using the TP in addition or as an alternative to the non-TP tools. Many non-TP examples were documented in Mabin and Balderstone (2000), while a recent example, Pirasteh and Farah (2006), documents a study combining TOC’s 5FS with Lean and Six Sigma with remarkable results. One wonders whether the results would be significantly different if the TP were used, rather than just the 5FS. Thus, further investigation relating to the methodological appropriateness of different combinations, or sequenced use, of TP tools in specific situations is desirable, as has also been suggested by Dettmer.

It may also be worth investigating whether the conventional sequenced use of TP tools should be followed “blindly.” While Dettmer (2007) promotes sequenced use of the TP tools, Schragenheim (1999) advocates a freer form of diagramming using the principles of the TP logic without confining it to specific diagrams. In addition, the TP tools can also be used individually to improve performance in a variety of situations, while many different combinations of

TP tools used in different orders have been found to be effective, as reported in the literature (see Tables 23-1 and 23-2). It could be helpful to identify the circumstances in which particular combinations or sequences may be most effective.

Ronen (2005) has issued the challenge to TOC researchers to confirm and improve TOC methods and apply academic rigor to TOC-related research and research on TOC. In this chapter, we have drawn on our classificatory examination of the philosophical underpinnings of the TOC TP and their relationship to different phases of problem solving (Davies et al., 2005) to show how such tools and methods purposefully attend to different issues and surface different insights, using different kinds of information sourced in different ways. We have shown how the choice and use of a TOC TP tool reflects, in essence, a deliberate attempt to represent, frame, or model a problem situation in a certain way, each representation being used with specific intent, thereby highlighting certain aspects while downplaying or ignoring other aspects. These matters are reflections not only of what the tool or method is intended to do, but what it assumes to exist—its *ontological* base—and the nature of what is represented or modeled, with what kind of information; that is, its *epistemology*. Consequently, we also see value in research that embraces such philosophical and methodological foundations to consider future developments of TOC methodology that may occur (1) via evolution of new tools, for example, new TOC TP; (2) via those tools that have yet to reach the peer-reviewed public domain, such as the S&T trees⁷; or (3) via the development of new application areas. Such research would need to embody the academic rigor necessary to build the academic stature of TOC.

We also see value in research that addresses shortcomings of the TP, as for example, the surfacing,

representation, and definition of feedback. In addition, research that targets new classes of problems and applications would be welcome, as would research to address matters of practicability and ease of use. Related research that seeks to aid reflection and learning about TOC methods-in-use, reasons for success or failure, etc. would prove useful for practitioners and underpin longitudinal work on the effectiveness of TOC tools and methods. Similarly, research that explores the psychological and technical barriers to the use of TOC TP tools would not only benefit practitioners, but also contribute to the development of strategies and resources for teaching TP in the *TOC for Education*⁸ program. Finally, given such an extensive agenda, there is a related need to coordinate such research efforts if they are to add to the TOC body of knowledge.

The classificatory mapping of the various TOC frames, models, and methods to the Mingers and M-B frameworks shows that they not only overlap or substitute for each other to some degree, in terms of purpose and underlying philosophical assumptions, but that they may also be complementary, not only in nature, but also in terms of insights generated about the problem. As stated elsewhere (Davies et al., 2005), the recommended pragmatic adoption of a multi-method or multi-methodological approach accords with the views of Burrell and Morgan (1979) and Brocklesby (1993) in their discussion and acceptance of the efficacy of multi-paradigm and multi-methodology development.

Our reviews of the TP literature show that seldom are the TP tools used in isolation. Certainly, for complex problems, there is evidence that several tools may be used as problem-solving intervention moves through diagnosis to the implementation phase. Such multi-method use is in keeping with the findings of analysis using the M-B classificatory framework, where we recognize that TOC methods can be used as complements to broaden or heighten the *appreciation* phase of intervention, or to

complement *analysis* and *assessment* or *evaluation* with a stronger *action* or *implementation* phase.

Indeed, when the full set of TOC tools and methods discussed are mapped to the M-B framework (see Tables 23-5 and 23-6), we note how TOC methods comprise a comprehensive multi-method approach, and can be regarded as a methodological set, a multi-methodological approach, a meta-methodology, or a meta-framework. We also see TOC and the TP tools as offering a complementarity that others have sought through the development of multi-method and multi-methodological approaches combining hard and *soft OR* methodologies and methods (Davies et al., 2005).

TOC can be described as a methodology that offers methods that embrace the whole range of activities or phases from problem identification and representation, the setting of appropriate objectives, generation and evaluation of alternatives, through to implementation. In forming this view, it has been instructive to surface and clarify the various activities embraced by TOC (see [Table 23-6](#)), as well as the nature of the philosophical assumptions, ontological and epistemological, that underpin the various methods and tools that make up TOC (see [Table 23-7](#)).

As previously noted, various authors have identified elements additional to the familiar TOC questions of *What to Change?*, *To What to Change?*, and *How to Change?* Research is needed to explore further all phases of problem-solving that contribute to improvement in organizations, to go beyond the *What to Change?*, *To What to Change?*, and *How to Change?* questions and phases, and to extend these questions to include and begin with *Why Change?* and to follow them with *How to Sustain the Change?* and *How to Establish a Process of Ongoing Improvement (POOGI)?* Articles defining these elements and logically connecting them as a system for improvement would be of value. These questions are, of course, preceded by questions relating to:

What the System is, What the System Goal is, and How Progress Toward the Goal will be Measured.

As such, our analysis has helped clarify the potential supplementary or complementary role of the TP tools in relation to traditional OR/MS methodologies and methods. In a general sense, we have commented on the seeming equivalence between TOC TP and *soft* OR methodologies like SSM. In particular, we have noted the equivalent roles filled by *rich pictures* within SSM and the CRT within Dettmer's (2003) broader use of the OODA process for strategy development. As such, there is much to be gained from reconceptualizing TOC and the TOC TP as being within the broader domain of problem-solving methodologies such as OR/MS, or within the specific domain of *soft* OR, not just as an academic discipline worthy of study, but as a meta-methodology that offers a set of methods for use alongside traditional OR/MS methods and other PSMs.

TOC methods have yet to be fully understood or endorsed by the OR/MS community. Similarly, we suggest that TOC methods have yet to be fully understood by the TOC community, in terms of their philosophical underpinning, their systemic nature as a multi-methodological set, and their multi-methodological use with other OR/MS and systems methodologies. The TOC community has yet to identify with the OR/MS and other kindred communities. Yet TOC embraces and can be embraced by OR/MS and *soft* OR. A next step is to continue to build awareness of such complementarity, and to understand more about how and when a multi-method approach can be used best.

As such, we see benefit in future research addressing multi-methodological issues, not just identifying the potential for combining methods in multi-method or multi-methodological use, not just in combining methodologies in multi-methodological use, but also assessing and clarifying the philosophical and methodological assumptions that would underpin methodological consistency and rigor in

using TP in harness with other methods and tools. For example, the notion of problem templates or archetypes is well founded and accepted in the systems world in terms of identifying common systemic structure in problematic situations using CLDs (Senge, 1990; Wolstenholme, 2004). Thus, there may be merit in exploring and developing archetype clouds for archetypical dilemmas, and in the development of archetypical solutions or solution processes.

Links to Other Chapters in the TP Section

The discussion in this chapter may usefully shed light on the nature of other TOC tools and methods, their use in multiple problem domains, and their potential for use in multi-methodological intervention. As a consequence, links to other chapters may prove fruitful in focusing attention on the purpose of design, the purpose of use of the TP tools, and the other philosophical assumptions that are made about cause-effect relations, how we surface them, and how we represent them in the particular forms that are manifest as the logic trees, a belief in the existence of root causes, etc.

In addition, having demonstrated the nature of the suite of TP logic tools as being a comprehensive methodology or meta-methodology, the classificatory frameworks used in doing so may be used to shed light on the efficacy of different TP tools used in combination with each other, or used in combination with other non-TOC tools or methods, or subsumed, for example, within the OODA process developed by Dettmer ([Chapter 19](#), this volume) to surface strategic issues and goals. Similarly, they may be used to shed light upon the S&T tree (as in [Chapters 15](#), [18](#), [22](#), [25](#), and [34](#) in this volume).

“Once you have solved someone’s problem, you have forever blocked them from inventing those answers for themselves.”—Goldratt (1990b, 18)

References

Abbott, A. 1988. *The System of Professions: An Essay on the Division of Expert Labor*. Chicago: University of Chicago Press.

Ackoff, R. L. 1977. “Optimization + objectivity = opt out,” *European Journal of Operational Research* 1(1):1-7.

Ackoff, R. L. 1978. *The Art of Problem Solving*. New York: Wiley.

Ackoff, R. L. 1979. “The future of operational research is past,” *Journal of the Operational Research Society* 30:93-104.

Altshuller, G. 1973. *The Innovation Algorithm* (Translated by L. Shulyak, S. Rodman 1999), Worcester, MA: Technical Innovation Centre Inc.

Anon. The Oxford Story, Goldratt Consulting Europe Ltd. <http://tocinternational.com/pdf/Oxford%20Radcliffe%20Hospital%20story.pdf>, retrieved 12 March 2010.

Antunes, J., Klippel, M., Koetz, A., and Lacerda, D. 2004. “Critical issues about the Theory of Constraints thinking process—A theoretical and practical approach,” *Proceedings of the 2nd World Conference on POM and the 15th Annual POM Conference*, April 30–May 3, Cancun, Mexico.

- Argyris, M. and Schön, D. 1974. *Theory in Practice. Increasing Professional Effectiveness*. San Francisco, CA: Jossey-Bass.
- Balderstone, S. J. 1999. "Increasing user confidence in system dynamics models through use of an established set of logic rules to enhance Forrester and Senge's validation tests," *Systems Thinking for the Next Millennium*, Wellington, VUW & the System Dynamics Society.
- Bazerman, M. 1996. *Judgement in Managerial Decision-Making*. New York: Wiley.
- Beer, S. 1985. *Diagnosing the System for Organisation*. Chichester: Wiley.
- Bennett, P. 1977. "Towards a theory of hypergames," *Omega* 5:749-751.
- Bohn, R. 2000. "Stop fighting fires," *Harvard Business Review* (July-August):83-91.
- Boyd, L. H. and Cox, J. F. 1997. "A cause-and-effect approach to analysing performance measures," *Production and Inventory Management Journal* 38(3):25-32.
- Boyd, L., Gupta, M., and Sussman, L. 2001. "A new approach to strategy formulation: Opening the black box," *Journal of Education for Business* 76(6):338-344.
- Brailsford, S., Harper, P., and Shaw, D. 2009. "Editorial—Milestones in operational research," *Journal of the Operational Research Society* 60(Supplement 1).

- Brocklesby, J. 1993. "Methodological complementarism or separate development—Examining the options for enhanced operational research," *Australian Journal of Management* (18)2: 133-158.
- Burrell, G. and Morgan, G. 1979. *Sociological Paradigms and Organisation Analysis: Elements of the Sociology of Corporate Life*. London: Heinemann Educational Books Ltd.
- Button, S. 1999. "Genesis of a communication current reality tree—the three-cloud process," *Constraints Management Symposium Proceedings*, March 22-23, Phoenix, AZ, 31-34.
- Button, S. 2000. "The three-cloud process and communication trees," *Constraints Management Symposium Proceedings*, March 13-14, Tampa, FL, 119-122.
- Checkland, P. and Scholes, J. 1990. *Soft Systems Methodology in Action*. Chichester: Wiley.
- Choe, K. and Herman, S. 2004. "Using Theory of Constraints tools to manage organizational change: A case study of Euripa Labs," *International Journal of Management & Organisational Behaviour* 8(6):540-558.
- Churchman, C. W. 1967. "Wicked problems," *Management Science* 14(4):B141-B142.
- Corbett, C. J. and van Wassenhove, L. N. 1993. "The natural drift: What happened to operations research?" *Operations Research* 41(4):625-640.

- Cox, J. F., Blackstone, J. H., and Schleier, J. G. 2003. *Managing Operations: A Focus on Excellence*. Great Barrington, MA: North River Press.
- Cox, J. F., Mabin, V. J., and Davies, J. 2005. "A case of personal productivity: Illustrating methodological developments in TOC." *Journal of Human Systems Management* 24:39-65.
- Cox, J. F. and Spencer, M. 1998. *The Constraints Management Handbook*. Boca Raton, FL: St. Lucie Press.
- Daellenbach, H. 1994. *Systems and Decision Making: A Management Science Approach*. Chichester: Wiley.
- Davies, J. and Mabin, V. J. 2007. "Investing in the research and science system—Government choices, systemic consequences," *International Journal of Business Strategy* VII(1): 56-71.
- Davies, J. and Mabin, V. J. 2009. "A systems perspective on the embedded nature of conflict: Understanding and extending the use of the TOC conflict resolution process using a multi-methodological approach," *The Systemist* 31(2&3):63-81.
- Davies, J., Mabin, V. J., and Balderstone, S. J. 2005. "The Theory of Constraints: A methodology apart?—A comparison with selected OR/MS methodologies," *Omega—The International Journal of Management Science* 33(6):506-524.
- Dettmer, H. W. 1995. "Quality and the Theory of Constraints," *Quality Progress* 28(4):77-81.

- Dettmer H. W. 1997. *Goldratt's Theory of Constraints: A Systems Approach to Continuous Improvement*. Milwaukee, WI: ASQ Quality Press.
- Dettmer, H. W. 1998. *Breaking the Constraints to World-Class Performance: A Senior Manager's/Executive's Guide to Business Improvement Through Constraint Management*. Milwaukee, WI: ASQ Quality Press.
- Dettmer H. W. 1999. "The conflict resolution diagram: Creating win-win solutions," *Quality Progress* 32(3):41.
- Dettmer, H. W. 2003. *Strategic Navigation: A Systems Approach to Business Strategy*. Milwaukee, WI: ASQ Quality Press.
- Dettmer, H. W. 2007. *The Logical Thinking Process: A Systems Approach to Complex Problem-Solving*. Milwaukee, WI: ASQ Quality Press.
- Doggett, M. 2004. "A statistical comparison of three root cause analysis tools," *Journal of Industrial Technology* 20(2):2-9.
- Doggett, M. 2005. "Root cause analysis: A framework for tool selection," *Quality Management Journal* 12(4):34-45.
- Eden, C., Jones, S., and Sims, D. 1983. *Messing About in Problems*. Oxford: Pergamon.
- Espejo, R. 2006. "What is systemic thinking?" *Systems Dynamics Review* 10(2-3):199-212.
- Foster, W. R. 2001. "And then there were nine layers of resistance," *Constraints Management Technical Conference Proceedings* 47-48.

- Friend, J. K. and Jessop, W. N. 1969. *Local Government and Strategic Choice: An Operational Research Approach to the Processes of Public Planning*. London: Tavistock Publications.
- Gass, S. 1989. "Model world: A model is a model is a model is a model," *Interfaces* 19(3): 58-60.
- Goldratt, E. M. 1990a. *The Haystack Syndrome: Sifting Information from the Data Ocean?* Crotonon-Hudson, NY: North River Press.
- Goldratt, E. M. 1990b. *What is This Thing Called The Theory of Constraints and How Should It Be Implemented?* Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 1994. *It's Not Luck*. Great Barrington, MA: North River Press.
- Goldratt, E. M. 1996. "Session 2: Giving Creative Criticism," *Managerial Skills Workshop*. New Haven, CT: Avraham Y. Goldratt Institute.
- Goldratt, E. M. and Cox, J. 1984. *The Goal*. Croton-on-Hudson, NY: North River Press.
- Goldratt, R. and Weiss, N. 2005. "Significant enhancement of academic achievement through application of the Theory of Constraints," *Human Systems Management* 24(1):13-19.
- Gupta, M., Boyd, L., and Sussman, L. 2004. "To better maps: A TOC primer for strategic planning," *Business Horizons* 47(2):5-26.

- Houle, D. T. and Burton-Houle, T. 1998. "Overcoming resistance to change—the TOC way," *1998 Constraints Management Symposium Proceedings*, April 16–17, Seattle, WA, 15–18.
- Howard, N. 1971. *Paradoxes of Rationality*. Cambridge, MA: MIT Press.
- Hrisak, D. M. 1995. "Breaking bottlenecks and TOC," *Chartered Accountants Journal of New Zealand* 74(7):75.
- Inman, R. A., Sale, M., and Green, K. W. 2009. "Analysis of the relationships among TOC use, TOC outcomes and organizational performance," *International Journal of Operations and Production Management* 29(4):341–356.
- Jackson, G. C., Stoltman, J. J., and Taylor, A. 1994. "Moving beyond trade-offs," *International Journal of Physical Distribution & Logistics Management* 24(1):4–10.
- Jackson, M. C. 2000. *Systems Approaches to Management*. New York: Kluwer Academic.
- Jackson, M. C. 2009. "Fifty years of systems thinking for management," *Journal of the Operational Research Society* 60(Supplement 1):S24–S32.
- Jamieson, N. R. 2007. *Breaking the Bottleneck! 10 Profitable Ways to Make the Theory of Constraints Work in Services*. London: ChangeNRJ Ltd.
- Kendall, G. I. 1998. *Securing the Future: Strategies for Exponential Growth Using the Theory of Constraints*. Boca Raton, FL: St. Lucie Press.

- Kim, S., Mabin, V. J., and Davies, J. 2008. "The Theory of Constraints thinking processes: Retrospect and prospect," *International Journal of Operations and Production Management* 28(2):155-184.
- Kim, W. C. and Mauborgne, R. 2005. *Blue Ocean Strategy*. Boston, MA: Harvard Business Press.
- Klein, D. and DeBruine, M. 1995. "A thinking process for establishing management policies," *Review of Business* 16(3):31-37.
- Kolb, D. A. 1984. *Experiential Learning. Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice Hall.
- Koljonen, E. L. and Reid, R. A. 1999. "Using system dynamics models to validate thinking process logic diagrams," *1999 Constraint Management Symposium Proceedings*, March 22-23, Phoenix, AZ, 67-76.
- Lockamy, A. and Cox, J. F. 1994. *Reengineering Performance Measurements: How to Align Systems to Improve Processes, Processes and Profit*. New York: Irwin Professional Publishing.
- Mabin, V. J. and Balderstone, S. J. 2000. *The World of the Theory of Constraints*. Boca Raton, FL: St. Lucie Press.
- Mabin, V. J. and Balderstone, S. J. 2003. "The performance of the Theory of Constraints methodology: Analysis and discussion of successful TOC applications," *International Journal of Operations and Production Management* 23(6):568-594.
- Mabin, V. J. and Davies, J. 2003. "A framework for understanding the complementary nature of TOC

frames: Insights from the product mix dilemma," *International Journal of Production Research* 41(4):661-680.

Mabin, V. J., Davies, J., and Cox, J. F. 2006. "Using the Theory of Constraints thinking processes to complement system dynamic's causal loop diagrams in developing fundamental solutions," *International Transactions in Operational Research* 13(1):33-57.

Mabin, V. J., Davies, J., and Kim, S. J. 2009. "Rethinking tradeoffs and OR/MS methodology," *Journal of the Operational Research Society* 60:1384-1395.

Mabin, V. J., Forgeson, S., and Green, L. 2001. "Harnessing resistance: Using the Theory of Constraints to assist change management," *Journal of European Industrial Training* 25(2/3/4):168-191.

Mabin, V. J. and Gibson, J. 1998. "Synergies from spreadsheet LP used with the Theory of Constraints—A case study," *Journal of the Operational Research Society* 49(9):918-927.

Mann, D. and Stratton, R. 2000. "Physical contradictions and evaporating clouds," *TRIZ Journal* April:1-12.

Mason, R. O. and Mitroff, I. I. 1981. *Challenging Strategic Planning Assumptions: Theory, Cases and Techniques*. New York: Wiley.

Mingers, J. 2000. "An idea ahead of its time: The history and development of soft systems methodology," *Systemic Practice and Action Research* 13(6):733-756.

Mingers, J. 2003. "A classification of the philosophical assumptions of management science methods,"

- Journal of the Operational Research Society* 54:559-570.
- Mingers, J. 2009a. "Taming hard problems with soft OR," *OR/MS Today* April 6(2):48-53.
- Mingers, J. 2009b. "The case for soft OR," *OR/MS Today* April 6(2):21-22.
- Mingers, J. and Brocklesby, J. 1997. "Multimethodology: Towards a framework for mixing methodologies," *Omega—International Journal of Management Science* 25(5):489-509.
- Moura, E. C. 1999. "TOC trees help TRIZ," *TRIZ Journal* 1-9.
- Munro, I. and Mingers, J. 2002. "The use of multimethodology in practice—Results of a survey of practitioners," *Journal of the Operational Research Society* 59(4):369-378.
- Noreen, E., Smith, D. A., and Mackey, J. T. 1995. *The Theory of Constraints and its Implications for Management Accounting*. Great Barrington, MA: North River Press.
- Nutt, P. C. 2002. *Why Decisions Fail—Avoiding the Blunders and Traps that Lead to Debacles*. San Francisco: Berrett-Koehler Publishers.
- Pidd, M. 1996. *Tools for Thinking: Modelling in Management Science*. Chichester: Wiley and Sons.
- Pirasteh, R. M. and Farah, K. S. 2006. "Continuous improvement trio," *APICS Magazine* May:31-33.

- Rahman, S-U. 1998. "TOC: A review of the philosophy and its applications," *The International Journal of Project Management* 18(4):336-355.
- Reid, R. A., Scoggin, J. M., and Segellhorst, R. 2002. "Applying the TOC thinking process: A case study," *2002 SIG Technical Conference Proceedings*, April 15-16, St. Louis, MO, 84-92.
- Reid, R. A. and Shoemaker, T. E. 2006. "Using the Theory of Constraints to focus organizational improvement efforts: Part 1—defining the problem," *American Water Works Association Journal* 98(7):63-75.
- Ricketts, J. A. 2008. *Reaching the Goal: How Managers Improve a Services Business Using Goldratt's Theory of Constraints*. Upper Saddle River, NJ: IBM Press, Prentice Hall-Pearson.
- Ritson, N. and Waterfield, N. 2005. "Managing change: The Theory of Constraints in the mental health service," *Strategic Change* 14(December):449-458.
- Rizzo, T. 2001. "TOC overview: The Theory of Constraints," *TOC Review* 1(1):12-14.
- Ronen, B. 2005. "Guest editorial: Special issue on the Theory of Constraints—Practice and research," *Human Systems Management* 24(1):1-2.
- Ronen, B., Pliskin, J. S., and Pass S. 2006. *Focused Operations Management for Health Services Organizations*. San Francisco, CA: Wiley/Jossey-Bass.
- Rosenhead, J. 1989. *Rational Analysis for a Problematic World*. Chichester: Wiley.

- Rosenhead, J. 2009. "Reflections on fifty years of operational research," *Journal of the Operational Research Society* 60(Supplement 1):S5-S15.
- Rosenhead, J., Elton, M., and Gupta, S. K. 1972. "Robustness and optimality as criteria for strategic decisions," *Journal of the Operational Research Society* 23(4):413-425.
- Royston, G. 2009. "One hundred years of operational research in Health—UK 1948—2048," *Journal of the Operational Research Society* 60(Supplement 1):S169-S179.
- Russo, J. and Schoemaker, P. J. H. 1989. *Decision Traps*. New York: Simon and Schuster.
- Scheinkopf, L. 1999. *Thinking for a Change: Putting the TOC Thinking Processes to Use*. Boca Raton, FL: St. Lucie Press.
- Schön, D. A. 1983. *The Reflective Practitioner. How Professionals Think in Action*, London: Temple Smith.
- Schrageheim, E. 1999. *Management Dilemmas*. Boca Raton, FL: St. Lucie Press.
- Schrageheim, E. and Dettmer, H. W. 2001. *Manufacturing at Warp Speed: Optimizing Supply Chain Financial Performance*. Boca Raton, FL: St. Lucie Press.
- Schrageheim, E. and Passal, A. 2005. "Learning from experience: A structured methodology based on TOC," *Human Systems Management* 24:95-104

- Senge, P. M. 1990. *The Fifth Discipline*. Sydney: Random House.
- Shoemaker, T. E. and Reid, R. A. 2006. "Using the Theory of Constraints to focus organizational improvement efforts: Part 2—Determining and implementing the solution," *American Water Works Association Journal* 98(8):83-96.
- Simchi-Levi, A. 2009. "Editorial comment," *OR/MS Today* April 2009, 21.
- Simon, H., Dantzig, G., et al. 1987. "Decision-making and problem-solving," *Interfaces* 17:11-31.
- Smith, M. and Pretorius, P. 2003. "Application of the TOC thinking processes to challenging assumptions of profit and cost centre performance measurement," *International Journal of Production Research* 41(4):819-828.
- Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>
- Tanner, J. F. and Honeycutt, E.D. 1996. "Reengineering using the Theory of Constraints: A case analysis of Moore Business Forms," *Industrial Marketing Management* 25:311-319.
- Taylor, L. J., Murphy, B., and Price, W. 2006. "Goldratt's thinking process applied to employee retention," *Business Process Management Journal* 12(5):646-670.
- Taylor, L. J. and Poyner, I. 2008. "Goldratt's thinking process applied to the problems associated with trained employee retention in a highly competitive

labor market," *Journal of European Industrial Training* 47(9):594-608.

Taylor, L. J. and Thomas, E. E. 2008. "Applying Goldratt's thinking process and the Theory of Constraints to the invoicing system of an oil and gas engineering consulting firm," *Performance Improvement* 47(9):26-35.

Thompson, N. 2003. "Best practice and context-driven building a bridge," *International Conference on Software Testing, Analysis & Review*, May 12-16, Orlando, FL.

Umble, M. and Umble, E. J. 2006. "Utilizing buffer management to improve performance in a healthcare environment," *European Journal of Operational Research* 174(2):1060-1075.

Umble, M., Umble, E., and Murakami, S. 2006. "Implementing TOC in a traditional Japanese manufacturing environment: The case of Hitachi Tool Engineering," *International Journal of Production Research* 44(10):1863-1880.

Watson, K. J., Blackstone, J. H., and Gardiner, S. C. 2007. "The evolution of a management philosophy: The theory of constraints," *Journal of Operations Management* 25:387-402.

Wolstenholme, E. 2004. "Using generic systems archetypes to support thinking and modeling," *System Dynamics Review* Winter, 20(4):341-356.

Wright, J. and King, R. 2006. *We All Fall Down: Goldratt's Theory of Constraints for Healthcare Systems*. Great Barrington, MA: North River Press.

Zeleny, M. 1981. "On the squandering of resources and profits via linear programming," *Interfaces* 11(5):101-107.

Zotov, D., Hunt, L., and Wright, A. C. 2004. "Analysing systemic failure with the Theory of Constraints," *Human Factors and Aerospace Safety* 4(4):321-354.

About the Authors

Victoria Mabin is Associate Dean (Teaching and Learning) in the Faculty of Commerce and Administration at Victoria University of Wellington, New Zealand, and Associate Professor in the Victoria Management School, where she teaches and researches a range of problem-solving and decision-making methods, specializing in TOC and hard and soft OR/MS methods. Prior to joining VUW, she worked for the New Zealand Government's scientific and industrial research organization, working as a consultant to business, government, and industry on a range of strategic and operational problems. Vicky graduated BSc (Hons) 1st Class from Canterbury University and PhD in Operational Research from the University of Lancaster, UK. She is a Jonah and holds the TOC ICO certifications in Thinking Process and Supply Chain Logistics, and serves on the examinations committee. She is a Fellow of the Operational Research Society (UK) and has held many positions with ORSNZ and NZPICS including President and Branch Chair, and as Editor and Editorial Board member for the *International Transactions in Operations Research*. She has published widely in books and international journals, co-authored *The World of the Theory of Constraints*, and given numerous academic and practitioner presentations and workshops.

John Davies is Professor of Management Studies and former Head of School, Victoria Management School, Victoria University of Wellington, New Zealand. He graduated from the University of Wales and the University of Lancaster with a background in operational research, and has developed his research interests primarily within the fields of the decision sciences, systems methodologies, and sports management. He is a Jonah and has published in leading academic journals spanning the decision sciences, technology management, systems, and sports management

He has been a council member of the Australian and New Zealand Academy of Management, Vice-President of the Operational Research Society of New Zealand, President of the Wellington Rugby Football Union, and is currently Vice President with Western Decision Science Institute.

Acknowledgments

We would like to acknowledge the valuable contribution of Hadley Smith with respect to data compilation and analysis in the latter stages of preparing this chapter. We would also like to acknowledge the sterling efforts of the editors: their insightful remarks, encouragement and support have played a key role in bringing this chapter into being.

CHAPTER 24

Daily Management with TOC

Oded Cohen

Introduction—Purpose of the Chapter

This book contains a blend of *Theory of Constraints* (TOC) methodology and standard solutions that have been developed, implemented, and perfected over almost three decades. This chapter is about giving managers the *Thinking Processes* (TP) tools and procedures to enhance their ability to make better decisions, implement them, and get the expected outcomes. To manage the TOC way we need a common basic agreement:

The role of managers is to ever improve the performance of the area under their responsibility.

Management responsibility is the smooth operation of their area today as well as in the future. Therefore, management must solve today's problems as well as initiate improvements for better performance in the future. Many managers think that there is a trade off between spending time (or money) sorting out today's burning issues or spending the time (or money) on improvement initiatives. That leads to managers just dealing with fires and not enough time devoted to system improvements. We would like to offer the use of the TOC TP for daily operations of helping to solve problems in such a way that is good for the short term as well as laying the foundation for the future.

Explicitly for this book, you—the manager—may be in one of three time phases (before, during, and after) with reference to a TOC solution:

1. In preparation for implementing TOC. Your area is run in a conventional way—in line with the company's views and/or your views.
2. In the process (project) of implementing a TOC solution. Bringing a new approach to your area may raise many issues, problems, and conflicts between the “old” way and the new way. You want to provide leadership and hence you must address these issues in a way that will move the implementation forward while ensuring the support and collaboration of the relevant people.
3. The TOC solution is an integral part of the way you run your part of the organization. As such, there is a need to work systematically on daily problems ensuring that the spirit of the solution and the TOC way of managing is kept. The TOC way means that we are committed to continuous improvement. Many of the TOC applications contain buffers and *Buffer Management* (BM). BM provides management with many incidents of disruption to the flow. These incidents provide opportunities for improvement and for that, we need effective tools for analysis and solution development.

Copyright © 2010 by Oded Cohen.

Three major TP are available for daily use: the *Evaporating Cloud* (Cloud), the *Negative Branch Reservation* (NBR), and the *Intermediate Objectives* (IO) Map.

The Cloud is the heart of the TOC methodology. It helps us to understand the problem and develop a breakthrough solution. Thereafter, we need the NBR in order to strengthen the solution and the IO Map in order to prepare the implementation plan. The chapter will follow this sequence. It is intended to show how to apply these tools in managing day-to-day operations, as the title of this chapter suggests. For more details of TP methods, see the other chapters in this Section VI.

Solving Daily Problems

During the course of the day, unexpected problems crop up to disrupt your concentration. Many times, you are unable to set them aside but must address them before moving on. Understanding problem structure and being able to frame a problem while surfacing the relevant facts helps address these problems effectively. A simple way to learn Clouds is to try them on these daily problems. Let us take a close look at the Cloud.

Problem Investigation and Solution Development—the Cloud

The objective of this section of the chapter is to enhance your ability as a manager to make better decisions and find better solutions in cases where conflicting options and views block such solutions. The better solutions are achieved through surfacing choices that resolve these conflicts (dilemmas) underlying the problem—using the Cloud method. The Cloud is a logical diagram that represents the problem through five boxes that are connected through the logic of cause and effect. The Cloud comprises three types of statements:

- Statements captured in the boxes A, B, C, D, and D'—presenting the most important entities helping to verbalize the conflict.
- The underlying assumptions—presenting the logical arguments supporting the cause-and-effect relationships between the entities written in the boxes (the logical connections are denoted through the use of the arrows on the diagram).
- Potential injections—new entities that when introduced into the reality of the problem can cause the conflict to disappear (this is why the solution is also called “Evaporating Clouds”). Please note that while theoretically there are potential injections to break every logical connection on the Cloud, it is unlikely that logical connections between A and B, or between A and C, need to be broken because by definition B and C are the necessary conditions to achieve A. If we feel we need to break these arrows, then it means that the Cloud is not the true representation of the conflict or dilemma.

See [Fig. 24-1](#) for the format of the Cloud, the assumptions, and the injections.

In this chapter, I will cover the use of the Cloud¹ as a stand-alone application of TP for daily problems, especially those that managers have with issues that in their eyes prevent them from performing their jobs better. People are promoted to managerial positions due to their capabilities and past performance. Managers are put in charge of areas (departments, projects, processes, etc.) and people, and hence are constantly bombarded by system and people problems. Not all of the problems are easy to solve. Many times managers feel that solutions they have come to are not the best they could have produced. If you have this feeling, then this chapter is for you.

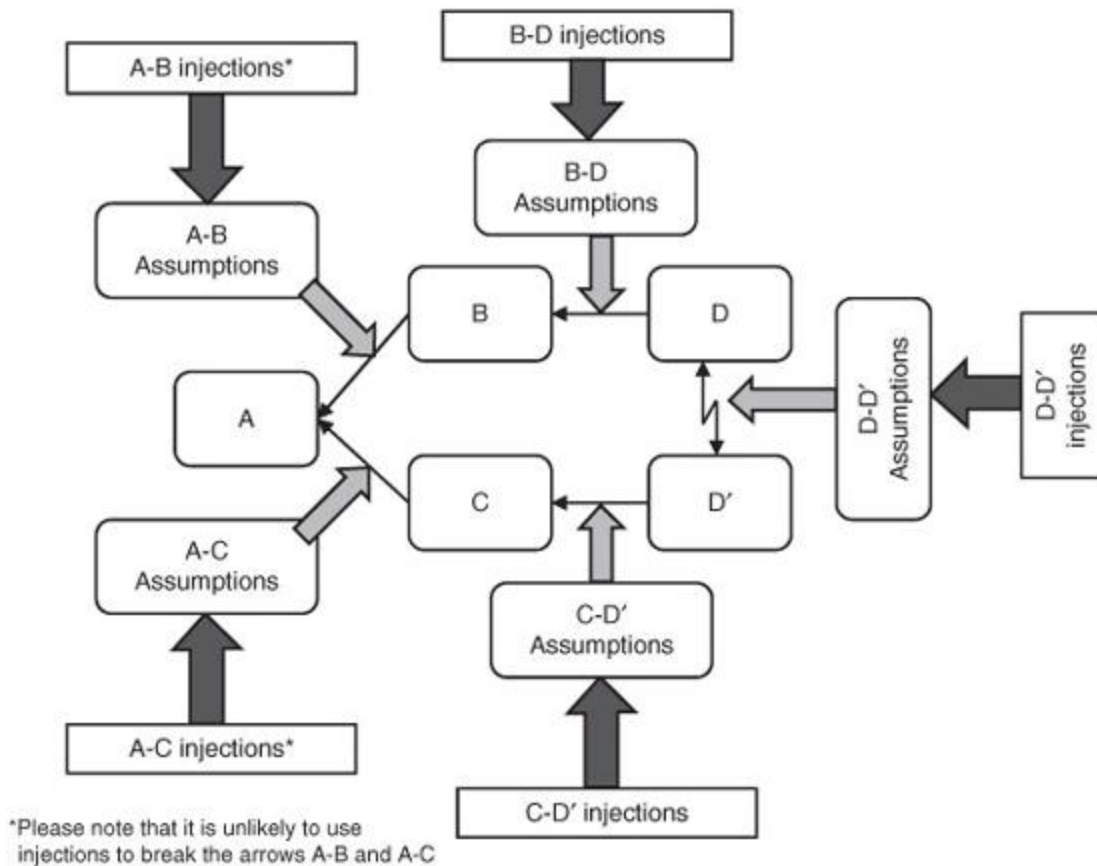


FIGURE 24-1 The general structure of a Cloud with the underlying assumptions and potential injections.

There is another good argument for using the knowledge and the tools of this chapter—to prepare you for the use of the TOC methodology for solving big issues that need the full TP work. We have found that people who have the knowledge and practical experience of the basic TP tools—the Cloud, the NBR, and the IO-Map—produce faster and better strategic solutions.

Let us start with solving daily problems.

When reviewing daily problems that managers encounter, we can see a broad spectrum of situations and challenges confronting managers while performing in their roles. On one side of the spectrum, they have to deal with their own inner dilemma of making a clear choice between options. On the other extreme, they have to deal with open conflicts

between them and other people in the organization or conflicts between two parties that they are expected to resolve. In between there are problems with the system or sporadically with other people (peers, supervisors, and even family members) that need to be addressed.

The objective of this chapter is to enhance your ability as a manager to address these problems in such a way that produces an immediate solution without blocking the long-term solution for these issues.

Application of the Cloud for Daily Problem Solving

Let us look at five applications of the Cloud for daily problem solving:

- Addressing **inner dilemmas**—issues when the person is faced with two major options and is not sure which route to take.
- Describing and solving **day-to-day conflicts** between two people.
- Analyzing **fire-fighting** situations—when the manager is forced to deal with emergency problems (fires) to find ways to prevent them from reoccurring in the future.
- Analyzing **a problematic area or a specific issue within the current reality**—by detecting an *Undesirable Effect* (UDE) in the area under analysis and building the UDE Cloud. The UDE Cloud is also instrumental for preparing for a sales meeting or developing a sales offer made better by understanding the reality of the buyer (the customer).
- Handling **multi-problem subjects** through the Three-Cloud approach—to help the manager build a more comprehensive view by building the **Consolidated** or **Generic Cloud** when there is more than one UDE. This approach is used for group consensus, accelerating

existing initiatives, and the buffer analysis for a *process of ongoing improvement* (POOGI).

All problems are handled by one general process of seven steps covering:

1. Building the Cloud and its logical components (Steps 1-5).
2. Constructing the solution (Step 6).
3. Communicating the solution to the relevant people (Step 7).

Building the Cloud is done through raising questions and writing the answers for each box. Thereafter, when you have a first version of the Cloud, apply the logical checks and make the necessary changes and upgrades.

The questions and the sequence of asking them differ from one Cloud application to another. The different applications of the Cloud also differ by the way we find a solution and the way we apply the solution and communicate it to the people who are involved and affected by the problem and the solution.

Let us start with addressing inner conflicts. Our experience shows that this type of problem is the easiest way for learning the mechanics of building the Cloud, as it does not pose any personal uneasiness in developing the solution and communicating it (we hope).

What Is a Cloud?

The Cloud² is the foundation in the TOC TP. It is, in my eyes, TOC in a nutshell. The Cloud is the process of framing the conflict and the generator of the breakthrough solutions. We use the term *breakthrough* in the sense that we bring to the

reality of the environment under study a new and fresh solution. Frequently solutions that were used under emergency conditions solved the problem but were not introduced into the system under the perception that they are not suitable for regular conditions of the system, many times due to perceived conflicts of the “emergency” solution with the current procedures of the existing system.

In TOC, we define something as a problem only if it prevents us from achieving what is important for us (our objective). Therefore, it is imperative to be able to verbalize what objective we are striving to achieve that is jeopardized by the problem.

At the same time, we know that if a manager complains that a problem cannot be solved there must be an underlying conflict that blocks him from finding and implementing solutions even though the objective that is being blocked is extremely important and the manager raising the problem has the interest and the desire to solve the problem. It feels like the third law of Newton—as if the manager applies force to solve the problem and experiences a “counterforce” that prevents him from sorting it out. The conflict is in the tactical level—actions or decisions that should be taken in order to achieve the desired objective.

Therefore, when a problem is brought under the heading of an “unsolvable” one, we need to reveal the underlying conflict by converting the problem into a Cloud. Once we have the Cloud, we can apply the problem-solving processes to reach a win-win solution. All Clouds have the same basic structure as that shown in [Fig. 24-1](#).

The Cloud is a five-box conflict diagram denoted by A, B, C, D, and D' (D prime). Each box has a specific role in describing the problem.

There are three different roles:

- Objective [box “A”]—the objective that is being blocked or jeopardized by this problem.
- Needs or necessary conditions [boxes “B” and “C”]—the term “need” is used in order to denote that this condition is *mandatory for* the achievement of the objective [A]. The $B \rightarrow A$ and $C \rightarrow A$ arrows present a logical connection of necessity. It reads, “In order to have the desired objective “A” we/I *must* have both needs B and C. The logic states that if one of these needs is missing, the objective will not be achieved.
- Tactics [boxes “D” and “D'”]—actions, wants, or decisions that are chosen to satisfy the needs. The $D \rightarrow B$ and $D' \rightarrow C$ arrows state that in order to satisfy the need, the specific action [D to satisfy the need B, and D' to satisfy the need C] *must* be taken. These actions, wants, or decisions cannot reside together at the same time and that brings them into conflict, which is denoted by the D-D' conflict arrow.

Ensuring the Quality of the Cloud—Logical Checks

As the Cloud is the base for finding a win-win solution, we have to ensure that it is built properly and that the logic is sound. After writing the Cloud, it is recommended that we read it again (even aloud) including the logical connections represented by the arrows:

In order to achieve “A” we/I ³ must have “B”

In order to achieve “A” we/I must have “C”

In order to satisfy “B,” action “D” must be taken

In order to satisfy “C,” action “D'” must be taken

D and D' are in direct conflict

Next, we should check the logic of the diagonals between the tactics or actions and the needs. The strong message from the Cloud is that every action endangers or jeopardizes

the achievement of the opposite need. The additional checks are:

- “D” endangers/jeopardizes/hurts need “C”
- “D” endangers/jeopardizes/hurts need “B”

After checking the logic, the necessary changes and upgrades should be made in order to make the Cloud clear and logically sound. It is also recommended to present the Cloud to a knowledgeable person who can give feedback about the clarity of the wording and the logic.

Solving Problems Using Clouds—the General Process

The general process is:

Step 1. Identify the type of the problem (inner dilemma, day-to-day conflict, etc.) and match it with the right type of Cloud to address such problems.

Step 2. Write a storyline of this problem in a factual, objective way as if you were completing an incident report. Objectivity is necessary even if the problem causes an emotional upset. The purpose is to unleash the intuition of the person building the Cloud about the problem and gather the data for building the Cloud.

Step 3. Build the Cloud.

Step 4. Check the logical statements of the Cloud and make necessary corrections and **upgrades**.

Step 5. Surface the assumptions behind the logical connections to find the one that is supporting the

conflict.

Step 6. Construct your solution and check it for win-win.

Step 7. Communicate the solution to the people involved in dealing with the problem.

Let us look at this process in detail for the example of an Inner Dilemma Cloud.

Inner Dilemmas

Step 1. Identify the type of problem and match it with the right type of Cloud to address such problems.

The *inner dilemma* is defined as a situation in which the manager is under pressure to take action or make a decision with which he or she doesn't feel comfortable. They have to choose between two conflicting options. They have not yet disclosed their preference, so there is no open conflict yet.

To learn and master the process of building and breaking the Cloud (finding the solution), a single problem is recommended. Avoid any problem addressing deep issues that contain a chronic problem⁴ or an unpleasant history of a relationship with someone else that may need a more comprehensive solution. An example of a single or "one-off problem" is as follows:

"I am under pressure from my boss to clear a technical request this Saturday, while I have promised my family a weekend out of town."

It is easier to learn the Cloud approach on such a problem rather than trying to deal with a problem of the chronic nature, like:

“My boss demands that I shall be instantly available for any work issues and often even on weekends and I cannot have any personal plans for the weekends.”

To demonstrate the process of the Inner Dilemma Cloud, I will use the following example of a single one-off problem:

“I am a project manager at an improvement initiative at a large hospital. I have resources assigned to the project; all of them also continue performing their daily jobs. One of them is not released by her boss to work on the project. What shall I do?”

Step 2. Write a storyline.

Write down in free format the facts about this, as if you were filing an official complaint or report. Explain in the report why this was a problem and how it has affected you or your performance.

Answer questions such as: Who—what—when—where? What did I want to do? Why? What did I feel forced to do? Why?

Example:

I am a project manager and Mary has been allocated as a resource to my project. She is from Bill’s department but Bill has no other involvement or responsibility for my project. He is not a sponsor and not a customer. I have assigned some tasks to Mary, which she has not done yet. When I have asked her why, she said that Bill prevented her from doing the tasks, as he didn’t agree with the approach we are taking. Mary has suggested that I go to Bill and sort this out, as Bill is very knowledgeable in the subject matter of my project. I don’t want to see him, but Mary is my friend. Mary seems to be between a rock and a hard place. My boss Fred is not willing to get involved and confront Bill.

Step 3. Build the Cloud.

The starting point of building the Inner Dilemma Cloud is the actions. We know what actions we are pressured to take, the ones with which we don't feel comfortable. We also know what actions we would prefer to take, but there is something that keeps us from explicitly taking these actions.

Hence, we have a good starting point on the Cloud—D and D'. From there we continue and build C and B and end up with A. Therefore, the sequence of building the Cloud is:

$$D/D' \rightarrow C \rightarrow B \rightarrow A$$

or alternatively,

$$D/D' \rightarrow B \rightarrow C \rightarrow A$$

Identifying D/D'

The idea here is to find out the major or most conflicting actions that one may consider in addressing this problem.

The guidelines are:

- Write down all the options that you have considered while trying to solve the problem.
- Split them into two groups: actions that you prefer to take and actions that you feel you are forced to take.
- Choose the one that you feel is the most distasteful or forced option and write it in the D box.
- Choose your preferred option and write it in box D'.

Example:

The list of tactics or actions that were considered and evaluated by the Project Manager is in [Table 24-1](#).

Considered Tactics/Action	Forced	Preferred
Ignore the whole thing.		X
Go to my boss—Fred.		X
Write Bill an email explaining our approach.	X	
Go and see Bill myself.	X	
Approach Bill in a committee that we both attend.		X
Tell Mary to go back to Bill.		X

TABLE 24-1 The Tactics/Actions Considered by the Project Manager

[D]: The most forced action: See Bill myself.

[D´]: The most preferred action: Ignore the whole thing.

Now we need to complete the Cloud.

After writing D and D´, you can move either to the B need or to the C need.

The sequence of moving to B or to C does not really matter. Some people find it easier to first write what it is they want rather than what is forced on them.

Write in box C the need that is satisfied by action D´ and check the logic: in order to achieve [C] I must [D´].

The project manager in our example wrote:

[C]: Get on with my work.

Check: In order to get on with my work—as a project manager—I must ignore the whole thing. I have more important things to do!

Write in box B the need that would be satisfied by taking the action in D and check the logic.

[B]: Fight for my resources

Check: In order to fight for my resources I must see Bill myself.

Well, this seems logical but is B verbalized as a need? For now, let us just go on with the process of building the Cloud and we will address that in the step of upgrading the Cloud. While learning how to build a Cloud it is important to proceed from “Good enough” criteria for the first version of the Cloud and not get stuck on one entity trying to figure out whether it is absolutely correct.

Write in box A the common objective that will be achieved by having need B and need C met (why are B and C so important? What for?) and check the logic.

The project manager wrote:

A: Able to deliver the project on time.

Check: In order to be able to deliver the project on time I MUST fight for resources and I MUST get on with my work.

The Cloud is a good enough Cloud for the next step of tightening its logic.

In summary, the sequence and the questions for building the Inner Dilemma Cloud are provided in [Table 24-2](#).

Box	Question to Guide in Writing the Content of the Box
D	What is the action or decision that I feel under heavy pressure to perform?
D'	What is the action or decision that I prefer the most?
C	What need (of mine) is satisfied by the most preferred action of D'?
B	What need (of mine) is satisfied by the most forced action of D?
A	What is the common objective that will be achieved by having need B and need C met?

TABLE 24-2 Sequence and the Questions for Building the Inner Dilemma Cloud

Step 4. Check the logical statements of the Cloud again and make necessary corrections and upgrades.

In Step 3, we write the entities in the boxes as the answers to the questions asked and check the logic of each arrow individually. In Step 4, we again check the logic of the

entire Cloud: $A \rightarrow B \rightarrow D$, $A \rightarrow C \rightarrow D'$ the conflict $D-D'$ and the diagonals D jeopardizing C and D' jeopardizing B .

Syntax Guidelines

Ensure that the entities in the boxes meet the following guidelines:

- Entities are whole sentences.
- Entities do not contain causality statements. Causality statements include words like if, because, sure to, in order to, etc.
- Entities D and D' are verbalized as actions and are in clear and direct conflict.
- Entities B and C are verbalized as clear and positive needs.

Let us check the example Cloud:

[A]: It is clear that the project manager cares about the project. She wants to do a good job. She is a capable and willing member of the hospital staff.

We can suggest A to be: [her objective is:] Deliver a successful project.

[B]: Fight for resources—is not verbalized as a need, it is an action (as it contains a verb indicating action—“fight”) that we take to satisfy the need “**having resources for the project** is necessary if we want to implement it.”

Therefore, we suggest upgrading the wording in B to: Have secured resources.

[D]: “See Bill myself” is one of the actions that will secure Mary as a resource for the project.

[C]: “Get on with my work” may explain the reasoning behind ignoring the whole thing—but it does not really work. “Get on with my work” is not a need. Here one has to be courageous and call a spade a spade. What can help us in finding a better C is the check on the diagonal—what does D jeopardize? From the text of the story line, we can derive that Bill’s attitude hurts the project manager’s feelings. Therefore, we can suggest need [C]: “Respect for my position as the chosen project manager.”
 [D’]: “Ignore the whole thing” is a decision.

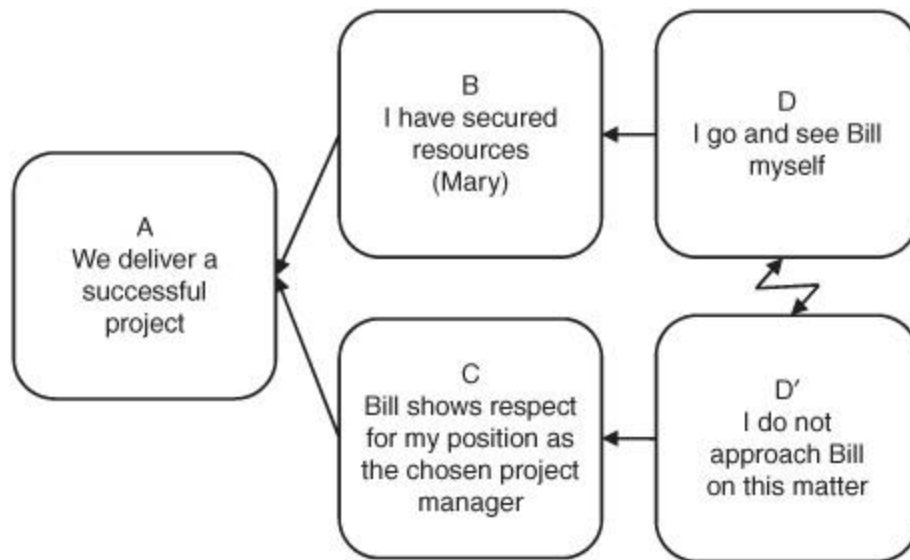


FIGURE 24-2 Example—the project manager’s Dilemma Cloud after the upgrading process.

Diagonal check: [D’] jeopardizes [B] because if this decision is taken it will jeopardize the need B as Bill will not release Mary to perform her project tasks unless he meets with the project manager. Yet, to clearly express that D-D’ are in a direct conflict we suggest to write in D’: “Do not approach Bill on this matter.”

Read the Cloud again to ensure it is logically sound. Now we have the upgraded Cloud as shown in [Fig. 24-2](#).

Step 5. Surface the assumptions that cause the conflicting tactics (actions and decisions). To better understand the conflict/dilemma and as a prerequisite for finding a solution, one should look for the reasoning behind the logical statements (of the arrows) and especially those that are leading to the conflicting entities of D and D'. The explanations behind the arrows state clearly why each of the boxes of the Cloud is absolutely necessary. In TOC terminology, we call them underlying assumptions. The way to surface them is by checking:

In order to have . . . (tip of the arrow), I must . . . (base of the arrow), because . . . Everything we state after the "because" is an assumption.

This is used for surfacing the horizontal arrows: $A \rightarrow B$, $A \rightarrow C$, $B \rightarrow D$, and $C \rightarrow D'$.

Please avoid repeating what is already stated by the existence of the arrow. Stating that in order to have C we must take action D' because the action D' is the only way to achieve C does not add any more understanding. Assumptions that explain only one part of the arrow also do not help the understanding. The assumption should establish the direct causal connection between the two parts of the arrow. Check that some of the words of the assumption refer to one box and some words refer to the other box.

Example:

$B \rightarrow D$:

In order for me (the project manager) to have secured resources (especially Mary) I must see Bill myself because . . .

[B-D 1]: Bill controls what Mary works on. This statement causally connects Bill to Mary.

[B-D 2]: Bill is blocking Mary from completing my project tasks. As far as I know, Bill does not release Mary from her

daily duties to perform tasks assigned to her according to our project plan.

[B-D 3]: Bill needs to be approached personally to get his collaboration on sharing his resources. This statement fits the syntax of an assumption as it explains the causal connection between B and D. It connects Bill and his “conditions” for releasing his people. Yet, the assumption is a bit one-sided and contains slightly negative views about the person involved. This statement can be reverbalized as: Bill usually wants to be consulted before releasing his resources.

C←D’:

In order to have respect for my position as the chosen project manager, I should *not* approach Bill because . . .

[C-D’ 1]: Yielding to local politics undermines my position.

[C-D’ 2]: Bill is not my boss, sponsor, or a customer of the project. (As Bill is not a part of the project community, approaching him will just weaken my position as the project manager and will be a loss of face.)

Surfacing assumption underlying D-D’:

These assumptions have to state clearly the reasons for the existence of the conflict. They have to explain why the two tactics stated in D and D’ are mutually exclusive and cannot coexist. They have to explain why the conflict cannot be resolved and what is causing the conflict to exist.

The statements that can help in surfacing the D-D’ assumptions are:

D and D’ are in conflict because . . . or,
I cannot resolve this conflict because

The logical arguments explaining the existence of the conflict can reveal different mind-sets, organizational

behaviors, policies, or procedures that drive opposing actions or decisions. It can reveal a shortage of something in common (like resources) and it can highlight a lack of mutual appreciation or confidence.

Example: D-D'—See Bill myself is in conflict with Do not approach Bill at all because . . .

[D-D' 1]: There is no procedure in the company that addresses a clash between a project assignment and routine departmental work of resources.

[D-D' 2]: I don't know what value there would be in meeting with Bill.

Graphically, the causality arrows are described in text boxes containing the underlying assumptions that are pointing to the relevant arrow as in [Fig. 24-1](#).

Step 6. Construct your solution and check it for win-win.

A solution to the problem is a change to the reality that removes a major reason for the existence of the Cloud. The way to achieve objective A is through the removing or invalidating one of the significant underlying assumptions. When a major assumption is invalidated, then there is no reason for the logical connection to exist and hence one of boxes may disappear from the reality, causing the conflict to disappear or evaporate. Hence, this process is called the *Evaporating Cloud* (EC).

Theoretically, we can challenge every arrow on the Cloud. However, from a practical point of view we want to make the changes on the tactical level; hence, we challenge the assumptions underlying $B \leftarrow D$, $C \leftarrow D'$ or $D'D'$.

Usually we can expect that after building the Cloud and the thorough logic check in Step 4 that the objective A and the needs B and C are well defined. B and C have been confirmed to be significant, positive, and necessary

conditions for achieving the objective stated in A. Accepting A, B, and C directs our efforts to solve the conflict between D and D'.

To invalidate or negate an assumption, we must introduce something new to replace it. This "something new" is called an injection.⁵

The injection is the change in reality that helps achieving the statement in the box at the tip of the arrow (B or C) without having the situation described in the box at the base of the arrow (D or D').

Under the current reality, the perception is that the only way to achieve B is through taking the action D and the only way to achieve C is through taking the action D'. The injection for B-D is a new facet to the reality so that B can be achieved. This injection is a valid solution only if it is also not in conflict with D'. Alternatively, the injection for C-D' is a new facet to the reality so that C can be achieved and B is not jeopardized. This injection is a valid solution only if it is also not in conflict with D.

Therefore, we have three potential options for breaking the Cloud:

- Injection to B-D assumption that replaces the use of D.
- Injection to C-D' assumption that replaces the use of D'.
- Injection to D-D' assumption that removes (or changes) both D and D' and suggests a new common tactic.

Conceptually, it is possible to find an injection for all three options.

Finding an injection is an important step in the process. The general recommendation is to try to think "outside the box" and ask yourself the following question: "In what situation is the stated connection between two boxes of the

Cloud not valid?” You may think about a different scenario, a different environment, or experience from the past in which the connection was not there. Many times we have the injection in our head under the heading of “I wish the situation was different . . .” We claim that when people have a problem bothering them they continuously think about solutions. Nevertheless, people tend to erase their own potential solution assuming that the solutions are unrealistic or impossible to implement. The EC method is the place to consider options and ideas that were dismissed before.

To practice the process of finding injections and to ensure that all possible options are considered, we recommend searching for injections to break the B-D, C-D', and D-D' arrows. Once we define a variety of injections, we can choose which one (or ones) we prefer to use. The selection of the arrow to break may have an impact of the acceptance of the suggested solution. The arrow, corresponding assumptions, and injections are shown in [Table 24-3](#).

Once you have all the potential injections, you can decide to choose one or more of them or even create a new injection that will take elements of some of the stated injections.

The solution has to work for you. That means that you are comfortable with it based on a better understanding of the problem and that you will feel happier when this injection becomes a part of your reality.

After choosing the injection, check the new reality in which your injection will replace one or both actions (D and/or D') and verify that the developed injection supports the achievement of both B and C.

Check: IF [injection] THEN I can achieve [B] and [C] without this conflict blocking me because . . .

If B and C are not achieved with the support of the injection, then it alone is not a good enough injection. This primary injection might need some supporting injections to

achieve both B and C. In some cases, rephrase the injection or select another one and check again.

Arrow	Assumption	Injection
B-D	Bill controls what Mary works on.	The Chief Executive directs Bill to release Mary to my project.
	Bill is blocking Mary from completing my project tasks.	Mary is transferred to work for me permanently.
C-D'	Bill is not my boss, nor a project sponsor or customer.	I accept that Bill is a part to my project (even though he doesn't have an official role).
	Going to Bill would be a loss of face.	I put the project needs in front of my personal feelings.
D-D'	There is a procedure for dealing with clashes of demands between project assignments and routine department work of resources.	I escalate the issue to the higher-level management to be addressed officially.
	I don't know what value there would be in meeting with Bill.	I carefully plan the meeting with Bill, making sure that he sees my needs and I listen to what he has to offer to my project.

TABLE 24-3 The Arrow, Corresponding Assumptions, and Injections for the Personal Dilemma

Please understand that if the solution were so simple, then you probably wouldn't be using the Cloud to identify the solution anyway.

If the injection replaces just one of the entities D or D', the check shall reveal that the injection fully replaces the removed action or decision. Once the actions to achieve B are not in conflict with the actions to achieve C, the Cloud disappears—it “evaporates!”

Scenario 1: If we break the B-D arrow it means that the injection replaces D, so the check has to be explicit: IF [injection] THEN B can be achieved and at the same time C will not be jeopardized. The reality after having the injection in place will be:

D' + Injection (breaking B-D)

Scenario 2: If we break the C-D' arrow, then the explicit check is: IF [injection] THEN C can be achieved and at the same time B will not be jeopardized. The reality will be:

D + Injection (breaking C-D').

Scenario 3: If we break the D-D' arrow, then it means that both D and D' are replaced by the injection and therefore we must check:

- IF [injection] THEN B can be achieved.
- IF [injection] THEN C can be achieved.

The three scenarios for breaking the Cloud create a new tactic that we can denote as D* (D star). The future after implementing the injection creates the diamond shape that replaces the Cloud, as shown in [Fig. 24-3](#).

Example—The D* injection as chosen by the project manager is:

I put the project needs in front of my own feelings;
listen to what Bill has to offer me as a project manager;
and negotiate resources (Mary) with him.

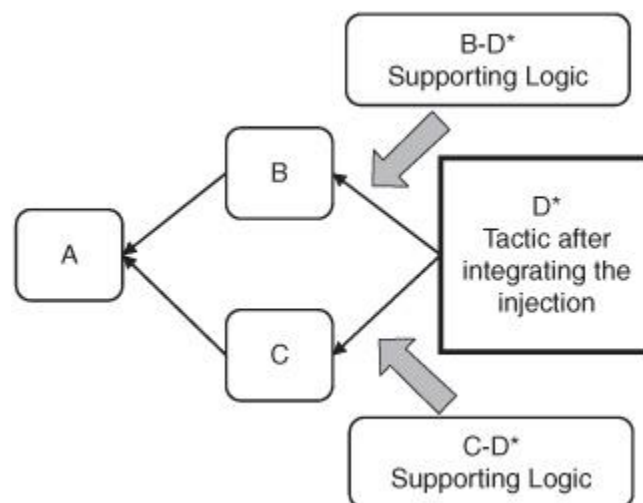


FIGURE 24-3 The diamond shape figure.

Check: If I do D^* , then I will have secured resources (Mary) because . . . approaching Bill the way he likes to operate is a good base for him meeting my need for resources for the project. (Bill is highly regarded in his professional area and is known to be tough but fair.)

If I do D^* , then I have increased my chances to have Bill's respect for my position as the chosen project manager because . . . showing respect to Bill increases the chances of him respecting me (based on Mary's recommendation to meet him as she probably knows Bill better than I do).

Summary of Step 6

That concludes the section of constructing a win-win solution for an inner dilemma using the Cloud method. A win-win solution the TOC way means that the tactics are not in conflict and that the solution supports both B and C needs. It means that we do not need to compromise on the achievement of the necessary conditions (B and C) and therefore we increase the chance of reaching the desired objective (A).

The learning experience covers:

- More choices to solve problems always exist than we think.
- A decision is the choice between conflicting options, often stemming from different mindsets and personal views.
- A problem is a blockage to progress caused by not resolving the conflicting tactics when these arise. Management cannot afford to procrastinate on making decisions as that leads to lose-lose situations.

- We should not give up on important needs—there must exist a non compromising solution. (As per the second basic concept of TOC—the existence of a win-win solution. The three basic TOC concepts are covered in the U-Shape section of this chapter.)
- A problem is not totally one person’s fault. Actually, in most cases we can reveal the system fault that causes the problem to happen.

In inner dilemma problems, both needs are the needs of the person facing the dilemma. When moving to the other types of Clouds we will find ourselves dealing with the needs of someone else—a person, an organizational function, or even the business needs. The definition of win-win stays the same—the achievement of both needs.

Step 7. Communicate the solution to the people involved.

We can define TOC as the ability to construct and communicate common sense solutions. Thus far, we have constructed the solution and now we have to consider and plan for implementing it.

In most cases, we need to achieve agreement, involvement, and support of the “other side” (the one who blocks us or conflicts with the actions or decisions we want to take). We define the injection as a win-win solution, but will the other side see the solution in that light? Therefore, we have to prepare carefully how to communicate the problem and the solution so that we get their agreement.

For the inner dilemma, the communication is simple. I just need to agree with myself that solving the problem is important and do what is necessary. Once the injection is in place and the benefit gained from sorting the problem, it reinforces our desire to use the tool more.

In the example, the communication is already an explicit part of the solution as the project manager is planning to have the meeting with Bill. Given that the issue is important and given that very thorough work has been done in developing the solution, it is better that the project manager plans her meeting with Bill.

The meeting has to be brief and focused on the desired outcome. The preparation should include the main points, the sequence, and some thought about potential pitfalls and questions with which she may be confronted.

When dealing with other problems, we will cover more aspects and more options for communicating the solution.

The Cloud is not only a technique; it is also a skill. We recommend practicing it regularly and frequently.

Day-to-Day Conflicts

Let's move to another very common type of problem—the day-to-day conflicts. These are conflicts between you and somebody else.

Recall the process outline:

Step 1: Identify the type of problem.

Step 2: Write the storyline.

Step 3: Build the Cloud.

Step 4: Check and upgrade the Cloud.

Step 5: Surface assumptions.

Step 6: Construct the solution.

Step 7: Communicate the solution.

Step 1: Identify the type of problem.

You can have different views than someone else; however, as long as you haven't clashed openly and publicly, you can handle the issue by using the Inner Dilemma Cloud. However, once the conflicting views are in the open, you have a bigger challenge.

To start with, we suggest you address simple conflicts, one-offs and not repeating problems. An employee is late to work is a one-off. However, when the same employee is late more than five times in the last two weeks, lateness starts to show a pattern of a repeating problem that poses an even bigger challenge.

In a day-to-day conflict, there are two definite sides —“Your” side and the “Other” side. The Cloud has a distinct structure as shown in [Fig. 24-4](#).

Reality provides us with many daily conflicts. It is not always possible to take a timeout in the middle of a disagreement or an open conflict in order to analyze the situation and develop a win-win solution. However, if the conflict has been concluded in a way that you find unsatisfactory, you may decide to take the time in the evening and deal with the problem using the Cloud method. The outcome of this effort can be, “Gee, I could have handled this problem better.”

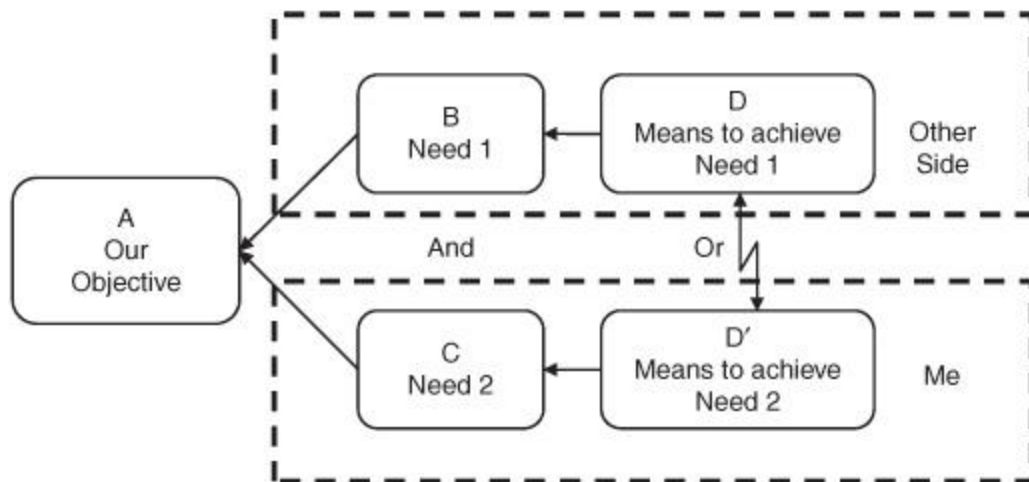


FIGURE 24-4 The general structure of the Day-to-Day Conflict Cloud.

Step 2: Write the storyline.

An example of a day-to-day conflict is described in the first page of *The Goal* (Goldratt and Cox, 1984)⁶:

When I finally get everyone calmed down enough to ask what's going on, I learn that Mr. Peach (the divisional vice-president) arrived at about an hour before, walked into my plant, and demanded to be shown the status of Customer Order Number 41427.

Well, as fate would have it, nobody happened to know about Customer Order Number 41427. So Peach had everybody stepping and fetching to chase down the story on it. And it turns out to be a fairly big order. Also a late one. So what else is new? Everything in this plant is late . . .

. . . As soon as he discovers 41427 is nowhere close to being shipped, Peach starts playing expeditor . . . Finally it's determined almost all the parts needed are ready and waiting—stacks of them. But they can't be assembled. One part of some subassembly is missing . . .

They find out the pieces for the missing subassembly are sitting over by one of the n/c machines, where they are waiting for their turn to be run. But when they go to that department, they find the machines are not setting up to run the part in question, but instead some other do-it-now job . . .

Peach does not give a damn about the other do-it-now job. All he cares about is getting 41427 out of the door. So he tells Dempsey (the supervisor) to direct his foreman, Ray, to instruct his master machinist to forget about the other super-hot gizmo and get ready to run the missing part for 41427. Whereupon the master machinist looks from Ray to Dempsey to Peach, throws down his wrench, and tells them they are crazy. It just

took him and his helper an hour and a half to set up for the other part that everyone needed so desperately . . .”

This is a day-to-day conflict. It is a one-off problem. Peach rarely visits the shop floor and does not tend to give instruction on how to run production. In this case, he cuts through the management hierarchy to give a direct instruction on what part to run on which machine. Yet when he does so, he gets into a conflict with the master machinist. This is an open conflict to the extent that the machinist throws his wrench and tells them they are crazy.

Step 3: Build the Cloud.

The starting point is the stated differences in the tactics D and D' (see [Fig. 24-4](#)).

For the sake of consistency, it is recommended to write in D from the viewpoint of the other side regarding the tactics—the actions or the decisions—and in D' from my view.

D and D' are different options and thus far I (or we) have not managed to come up with a workable compromise that will bridge between the two options.

In the example, the Machinist's side is C-D' and Peach's side is B-D.

The sequence of building the Conflict Cloud is as follows:

We start building the Cloud by stating D and D'. We can start with D or with D'.

In the example, the whole incident starts because Peach gives a direct instruction—hence, stating the [D] of the Cloud:

[D]: The tactic (action/decision) the other side (Peach) wants to employ.

[D:] Reset the machine to work on the missing part for order 41427 now.

[D']: The tactic (action/decision) I (the master machinist) want to take.

[D´]: Stick to the current setting to produce the other urgent part now.

[C]: The need I (the master machinist) am trying to satisfy or achieve by taking the tactic D´.

This follows the same way we have done it in the Inner Dilemma Cloud.

Once the conflict is clear, it is easier to move to [C]—my need—as the person who builds the Cloud is emotionally involved and has a clear view of why he or she is right in this conflict.

I am the Master Machinist. My job is to prepare the machines and get them ready for the jobs that need to be run. I want to do a good job and I want my work to be appreciated. Therefore, my need may be verbalized as:
[C]: Be acknowledged for my contribution to the production plan.

[B]: The need that the other side (according to my perception) wants to satisfy or achieve.

Many times, it is difficult to write the need B because when there was a heated discussion with the other side we were not attentive in listening to their arguments, and therefore we do not have a recollection of why the tactics they suggest (or demand) are important. Over time, with practice and experience, we will know how to identify a conflict and listen carefully to what the other side says so we can write the Cloud better. If we are not sure what to write in the B box, we may speculate under two conditions:

1. We write a need in a positive way.
2. If, during the Cloud communication, the other side corrects us and verbalizes their need, then we make the necessary corrections on our Cloud.

In this incident, the need of Peach is very clear as he cares only about one order. Customer Order Number 41427 is an order of Bucky Burnside—the biggest customer of the plant whom the management does not want to upset as they may lose him. Later in the chapter Peach tells Alex Rogo about the unpleasant telephone call he got from Bucky the night before. Of course, the machinist was not aware of that in the heat of the moment.

When in the conflict, people tend not to state their arguments, or if they do so, the other side is not always listening or recording these arguments.

Nevertheless, for building the Cloud the machinist must write his perception of what is the need that Peach was trying to achieve when instructing him to reset the machine. This is the most challenging part in building the conflict Cloud. We could write in *[B] not to upset the important customer of Order Number 41427*, but we want [B] to be worded in a positive way. Therefore, we have to ask ourselves another question—why is it so important not to upset this customer? This can be answered with *[B] secure the business with the important customer of Order 41427*.

[B]: Secure the business with an important customer

[A]: The common objective that we—the other side and me—collectively try to achieve. This is a tricky box. Usually the tactic of the other side blocks me or causes damage to my need and hence I do not see the common ground or collective objective. In the working environment, I may have conflicts with my subordinate, my boss, my peer, or an external person such as a vendor, service person, etc. We can find A by asking a simple question—why are we discussing this issue? Why are we in the same room?

The machinist knows that it is important to fulfill all the orders in time. He knows that this supports the financial

performance of the plant. Therefore, we can assume that both want the plant to be successful. In order for the plant to be profitable, all the orders must be fulfilled on time. Hence, we can conclude that both have a common objective: *[A]: Have a profitable plant now and in the future.*

See [Fig. 24-5](#) for the Cloud.

In summary, the sequence and the questions for building the Day-to-Day Conflict Cloud are provided in [Table 24-4](#).

Step 4: Check and upgrade the Cloud.

Follow the same process as for the dilemma Cloud.

Step 5: Surface assumptions.

Follow the same process as for the dilemma Cloud.

Step 6: Construct the solution.

In constructing the solution, we proceed from the assumptions that were surfaced in Step 5 to the injections. We end up with a list of potential injections. However, the situation may influence the choice of an injection to be the solution.

From one side, we are driven by the desire to move on with the action we want to take (D') as this will help us in achieving our need, and hence the tendency is to push and persuade the other side to see our point of view and accept that our D' is the right one!

From the other side, this approach will hardly work as we have already tried it and we have failed to convince the other side.

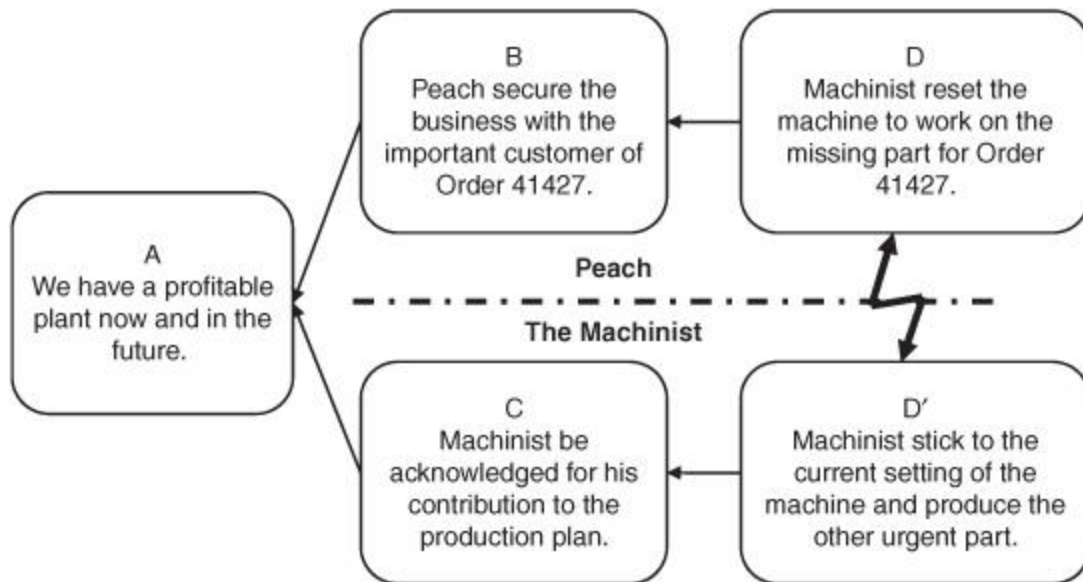


FIGURE 24-5 An example of the Day-to-Day Conflict Cloud: conflict between the master machinist and Peach.

Box	Question to Guide in Writing the Content of the Box
D	What is the tactic (action/decision) the other side wants to employ?
D'	What is the tactic (action /decision) I want to take?
C	What is the need I am trying to satisfy or achieve by taking the tactic D'?
B	What need does the other side (according to my perception) want to satisfy or achieve?
A	What is the common objective that we—the other side and me—collectively try to achieve by having need B and need C met?

TABLE 24-4 Sequence and the Questions for Building the Day-to-Day Conflict Cloud

Hence, we may want to employ a different approach. We break the Cloud on our side between C and D'! We shall find an injection that supports the achievement of our need C that can coexist with the D tactic—the tactic, want, or action of the other side.

Now the solution D^* is comprised of the D of the other side plus an injection that breaks the $C-D'$ connection. Usually, it is within our ability to perform the injection and hence the problem can be solved easily and should not be too difficult for the other side to accept.

In *The Goal*, this conflict was not resolved in a win-win way. The short-term need was stronger and the machine was reset, but the need of the machinist was not addressed. From the machinist's point of view, this was yet another example of management making crazy decisions. Can we find a win-win solution?

Given that the situation was critical and assuming that the machinist wanted to find a win-win solution (after the event), the focus should be on breaking the $C-D'$ connection.

Any assumption underlying $C-D'$ has to explain the logical reasoning between the two entities. As C is a positive and acceptable need, we have to understand why D' (not resetting the machine) is perceived by the machinist in this situation as the only way to achieve C (be acknowledged for the contribution).

One explanation can be that the machinist has just completed a long setup process (several hours). During the setup time, *no* production was done. This is pure downtime for a critical machine. By telling me [the machinist] to reset the machine, "they" (my managers) clearly radiate to me that my efforts were useless and not needed. I do not feel appreciated.

The assumption is that appreciation is measured by the efforts we put in. This assumption can be challenged. A potential injection can be: In this critical situation, management needs my support and willingness to make an extra effort and reset the machine again.

Is it a win-win or just a nice name for a compromise? Cynical people may say, "You have ended up doing what you were told in the first place!" We say it is a step in the

right direction. The conflict is driven by our emotions and our emotions are influenced by perceptions. It is correct that in this case to the outside world it looks like a compromise, but for the person addressing this problem it may bring a relief.

The major lesson that can be learned from this experience is that an open conflict is not that easy to resolve. Hence, maybe the next time the person who knows the Cloud method may control the reaction before the situation deteriorates.

Please note that using this approach too often with the same person—breaking the Cloud on your side—can create a situation in which the other side will expect you to always break the Cloud on your side. In the end, you would like these people to participate in solving the problems by breaking the Clouds on their side—after you have demonstrated your openness in dealing with problems and a willingness to “give up” on your initial want.

Step 7: Communicate the solution.

Generally, we recommend that once you find yourself in a “tug-of-war” situation, you should suggest taking time to think about the problem. If this is accepted, then you have the chance to build the Cloud, develop a solution, and then come back and communicate it.

If this is not the case and you or the other person imposed a solution, there is still value in using the analysis as a learning case to handle such situations better in the future.

In the example of the machinist, the resetting of the machine was imposed and done. The value is in addressing the emotion and considering the outcome if the injection suggested in Step 6 would have been used. In this case, the machinist should use the injection as a mode of operation for himself. Hence, he does not need to communicate the solution to anybody. Yet the managers of the machinist can benefit from such a mindset. They are not the “other side”;

they suffer the same problems and having another personal conflict does not help them at all.

In planning the communication, we use the TOC understanding of the layers of buy-in (decision making). They stem from the work that was done in investigating the resistance to change.⁷ This understanding recognizes that accepting a solution contains several layers. (I refer here to the views that contain five layers. Other TOC practitioners may use different numbers.) In preparation for the communication, we must cover all the layers that are relevant for addressing this problem. For a day-to-day conflict, we should prepare ourselves for the first three layers:

Layer 1—achieving agreement on the problem.

Layer 2—achieving agreement on the direction of the solution.

Layer 3—achieving agreement on the solution (that it will bring the desired outcome).

The other two layers will be introduced after the section of NBRs and obstacles:

Layer 4—achieving agreement that there are no negative effects.

Layer 5—achieving agreement that we can overcome implementation obstacles.

The problem is presented by the Cloud.

The direction is accepting the other side's tactic and breaking the Cloud on my side.

The desired outcome is checked in Step 6—the injection that breaks the C-D' side by replacing D' supports the achievement of C and does not hurt the achievement of B.

In the actual communication, we do not have to systematically follow the flow of the layers. We have to be

flexible to suit the preference of the other side. The process should be based on a face-to-face meeting. It is not recommended to try to sort out such problems using email.

Come to the meeting and say:

We have a difference of opinions on the issue of . . .

I have been thinking about it and I would like to work with you on finding a workable solution.

You want D and I want D'. These two are not compatible. I suggest we go with your D but we need to ensure that my C is taken care of as well. Do you have any suggestion how we can take care of it?

If the other side comes with any suggestions, you can check them against the potential injections that you have in mind to break the Cloud. If the suggested idea is close enough, you can agree on the solution. If not, continue in search for an amicable solution. Given that you are willing to contemplate suggested injections to your side of the Cloud, this discussion should be amicable and should end on a positive note⁸.

Reducing Fire Fighting

Managers have a huge impact on the performance of their systems. They need time and stamina to deal with improvements. Their time should be exploited. The opposite of exploitation is waste. One of the common causes for waste of time and disruption to the managerial process is known as fire fighting. In this section, we show how to use the Cloud method to address fire fighting and to improve the system to prevent such fires from reoccurring.

Step 1: Identify the type of problem.

Fire fighting is a common headache for managers. Irrespective of their own plans or issues that they are

expected to attend to, they are confronted by a sudden unexpected problem—a fire—that they are expected to solve immediately.

The manager is sitting in his office and then comes a knock at the door. A lieutenant (someone who reports to the manager) enters and says, “Boss, we have a problem.” What he really means is, “Boss, there is a problem and I need you to sort it out now, otherwise something unpleasant will happen to your area of responsibility.”

The nature of such a problem is that the boss has to stop everything and sort out the problem. Fire-fighting problems cause managers to do jobs and tasks that were supposed to be taken care of by their subordinates or not supposed to happen. That leads to the loss of valuable time and energy of management.

Addressing the fire-fighting problems in a systematic way—using the Cloud method—helps the manager to become more effective, less interrupted, and helps him to upgrade his subordinates’ skills and run his area more effectively.

Please note that the name may be a bit misleading. The method that we are about to suggest is not how to solve the fire-fighting problem itself. Fires are happening and the manager must find an immediate way to put the fires out. The idea is to use the incident of the fire for finding ways to prevent them from happening repeatedly. Therefore, a post-event analysis of the conflict is a conceptual one—what actions could have prevented the need of the manager to be sorting out the situation? Why couldn’t the people affected by the fire put it out themselves? When they cannot put out the fire, they come to their bosses and ask for help. This causes a disruption to the manager—the managerial fire.

Managerial fires are caused by three major reasons:

1. Lack of knowledge—the subordinates don’t know how to function in certain situations. The knowledge

resides with the manager or it is within his or her reach. This is caused by the manager's lack of time to transfer the knowledge and knowhow to their lieutenants and, in some cases, by lack of willingness of the lieutenants to learn from their bosses.

2. Lack of authority—the lieutenant has the responsibility but lacks the authority. Many times the system restricts the level of authority through tight policies and procedures.
3. Lack of confidence—the people who are supposed to take the necessary actions feel incapable of performing them and come to their managers to do the actions on their behalf. Many times people fear that they will be punished if something goes wrong while they are trying to solve a burning problem for the company.

The nature of the fire-fighting problem is that the problem must be attended to once it is raised. The manager must come up with an immediate solution and put out the fire. However, we have to approach the fire-fighting problem with the view that it is a manifestation of a system failure. It is beneficial to investigate the causes of the fire and take initiatives to remove them. Therefore, after the fire is dealt with, it is recommended that the manager examine the fire using the Cloud method to develop a solution to prevent this fire from reoccurring in the future by addressing the cause of the fire.

Step 2: Write a storyline.

Example:

I am the Customer Service Manager. Yesterday the person responsible for shipping products, who reports to me, came and asked for my help. There was a shipment for a particular customer that was due to ship but the

delivery location was not clear. The Customer Account Manager for this particular customer, who also reports to me, had been unavailable for 3 days.

I had to call the customer several times, and after some hassle, I got the information and gave it to the Shipping Clerk. Then I got back to my other work.

In this incident, the Shipping Clerk wants to perform his job properly as expected. We can assume that in the past when orders were shipped late this person was confronted and challenged—even when it was not his fault. He could have shrugged his shoulders and done nothing until the account manager came back. However, the Shipping Clerk cares! Hence, he goes to the Customer Service Manager and informs him about the problem.

The manager solved the problem by calling the customer himself. The customer was unhappy about the call. This problem was solved for now, but there is nothing to prevent the same problem from happening again in the future. This can be a good reason for the manager to investigate it using the Cloud method.

Step 3: Build the Cloud.

The sequence for building the Fire-Fighting Cloud is different from the previous two types because the trigger for the Cloud is different. For the Inner Dilemma Cloud and the Day-to-Day Conflict Cloud, the problem itself appears on the Cloud. The conflict is between two different tactics: D and D'. Once we write them, we can proceed to the B and C and eventually A.

In the fire-fighting scenario, the problem triggers the Cloud but is not recorded on the Cloud itself. We deal with the problem because it is extremely important. That means that the problem is jeopardizing the objective A and especially one of the needs. Therefore, the entry point to the Cloud is the need that is endangered. From there we

continue to fill up the boxes according to a logical flow and questions in [Table 24-4](#).

Example of building the Cloud of the shipping problem as seen by the Customer Services Manager:

[B]: Jeopardized need: Secure on-time shipment to customer. If we do nothing, the shipping details will not be obtained before the Account Manager is back and by then it will be too late.

[D]: Action to achieve B: The Shipping Clerk is allowed to call the customer for shipping details.

The suggested tactic in D is not too bad, but is not allowed according to the procedure stated in box D' and it may cause the negative implications that the procedure tries to prevent.

[D']: The blocking procedure: Only the Customer Account Manager makes all the calls to the customer.

[C]: The need that is taken care of by the procedure in D': Maintain good customer relationships.

Customers do not like having many people from their supplying companies call them and deal with different aspects of the products or services customers purchase from them. Therefore, we expect that if the Shipping Clerk calls the customer, he or she will be annoyed (as indicated by the storyline).

[A]: The objective: High level of customer service.

The Cloud is presented in [Fig. 24-6](#).

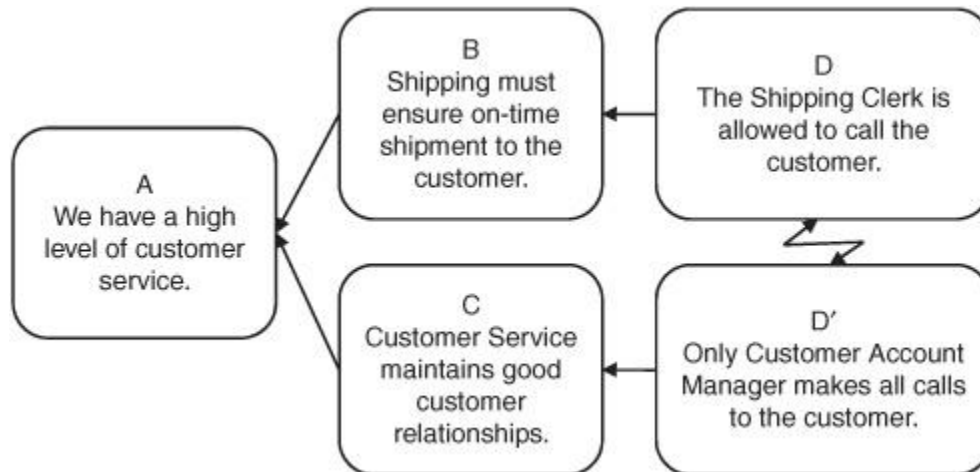


FIGURE 24-6 An example of a Fire-Fighting Cloud.

In summary, the sequence and the questions for building the Fire-Fighting Cloud are shown in [Table 24-5](#).

Step 4: Check the logical connections and upgrade.

Conduct the regular logical checks of the Cloud arrows.

Check that this problem puts you—the manager—in direct conflict with your system (sometimes even the one that you have put in place).

Check that the action in D jeopardizes C.

Check that the procedure or tactic of D' jeopardizes B.

Upgrade the Cloud.

Step 5: Surface assumptions.

Raise the relevant assumptions for all the arrows of the Cloud.

The assumption underlying A-B and A-C are needed to re-establish the importance of both needs. They are necessary conditions for the high level of performance of the area under the manager's responsibility. The assumptions should support the manager's intuition of why he has to deal with the fires.

Box	Question to Guide in Writing the Content of the Box
B	What important need of the system ^a does the fire jeopardize or endanger?
D	What action can be taken to meet the jeopardized need in B?
D'	What action or procedure is in place that prevents taking the actions suggested in D?
C	What other important need of the system demands the procedure that is stated in D'?
A	What is the common objective achieved with both B and C?

^aPlease note that we use the term "system" in this case to denote something which is beyond the need of the person raising the problem, and it must be within the area of the responsibility of the manager that is handling the fire-fighting problem.

TABLE 24-5 Sequence and the Questions for Building the Fire-Fighting Cloud

The C-D' assumptions support the reasoning for the system that was put in place to achieve C. The current procedures support the smooth running of the organization. Procedures are there to ensure the quality, consistency, and effectiveness of processes within the company. Therefore, the majority of the assumptions are strong and positive. However, as good and comprehensive as the procedures are, they do not cover the full spectrum of possible situations.

Hence, B-D assumptions reveal situations in which the existing procedures are weak.

^aPlease note that we use the term "system" in this case to denote something that is beyond the need of the person (the Shipping Clerk in this instance) raising the problem and it must be within the area of the responsibility of the manager (Customer Service) who is handling the fire-fighting problem.

D-D' assumptions reveal the reasons for the conflict. Usually they point at the rigidity of the procedures to the extent that they cause harm to business needs and toward

the lack of coverage in the procedures for dealing with emergencies and special situations.

Example:

C-D' Assumptions:

In order to [C] Maintain good customer relationships, all people of the Customer Service Department (including the Shipping Clerk) MUST follow the procedure that D' only the Customer Account Manager makes all calls to the customer BECAUSE. . . .

- We have a policy of “single point of contact” [just stating the source of the employed procedure].
- The customer gets irritated and confused when contacted by different people from the company [explains the logic that has brought the company to have this procedure].
- Anyone other than the Customer Account Manager will confuse the customer.

B-D Assumption:

- The Customer Account Manager may be out of reach and the shipping instructions are not clear. (This is a special case where the two conditions happen at the same time. If only one of these conditions would happen, then there would be no problem.)

D-D' Assumptions:

- The procedure clearly states that in any situation “No one is allowed to contact the customer but the Customer Account Manager.”
- The procedure does not cover rare situations like the one that has caused this fire.

Step 6: Construct a solution.

The purpose of using the Cloud method for the fire-fighting problems is the desire to exploit the management time better by removing the disruption caused by these problems. Therefore, we want to find a good and permanent solution to the problem, and for that we need to decide which arrow is better to break.

Ideally, we do not want to choose a solution that will violate the procedure. If we break the C-D', this means throwing out the existing procedure.⁹ This will be like “throwing the baby out with the bath water.” The solution has to be a combined injection that addresses three arrows at the same time: B-D, C-D', and D-D'.

In the example, the Customer Service Manager was forced to put out the fire. This is the first priority. In such situations, it is too late to find a win-win solution. The manager has to make a decision on the spot. In the example, the manager decided to call the customer himself. As such, B was salvaged but the customer may have been unhappy, which means that C could have been endangered. We do not want to pass judgment on the manager's decision, we just want to learn from it what can be done systematically to prevent it from reoccurring. If we allow someone other than the Customer Account Manager to call, it means we break C-D'. If we do not allow the Shipping Clerk to always call, then we break B-D.

If one action or one injection breaks both B-D and C-D', then it means that the injection also breaks at least one of

the assumptions underlying D-D'; in this case, it breaks the assumption that the procedure is rigid and must be adhered to no matter what the circumstances. The injection also challenges the perception that the procedure is comprehensive and covers all possible scenarios.

The direction for addressing fire-fighting situations (using the Cloud method) is to integrate the emergency solutions into the existing procedures. This direction stems from the appreciation that good ideas are invented and used in emergencies but are not accepted for regular times. As such, these ideas stay as the "assets" of the individuals and do not become a part of the organization's expertise.

Conclusion: Develop the solution by examining the actions used in emergencies, upgrade and formalize the actions so that they support both needs of the Cloud, and integrate them into the existing procedures.

In the example, the injection is the suggested amendment to the procedure:

Whenever on-time shipment is at risk from inadequate delivery location information and the Customer Account Manager is not available, then the Shipping Clerk has the authority to contact the customer about this information.

Check for win-win:

This amendment sorts out the specific situation of this fire problem. The Shipping Clerk can call the customer and the details are obtained. The customer will get the order on time. B is protected. But what about C? What about the assumption that the customer will be irritated if someone other than the Customer Account Manager approaches them? For that, we have to ensure that the customer is made aware **in advance** about this change to the procedure through explaining that this action will be done with the view of protecting

the interest of the customer and will be used only in rare cases. The customer should accept that.

Step 7: Communicate the solution.

There are two major steps in communicating the solution—getting the consensus with the relevant people and formally making the amendments to the procedures. The second part has to be done according to the company's way of making changes to the procedures. Let us start with getting the consensus. There are at least three parties involved in the solution for the fire-fighting case:

1. The manager who is prompted to put out the fire—the manager has the desire and stamina to sort out the problem as he or she is interrupted by the problem and suffers the consequences. As such, the manager should look for the views of the other parties and incorporate them in the solution while continuing to check that the solution is win-win.
2. The person/function that raises the problem—they want to perform their jobs in a way that will acknowledge their contribution. They are blocked by a lack of authority. They do not necessarily demand more authority, but they do not want to experience the negative consequences of their inability to perform their job.
3. The person/function that represents the need that triggered the procedure—they are generally in favor of the existing procedure as it supports the objective or deliverables of their jobs. Therefore, they may not be that happy to incorporate changes to the procedures.

Here is a suggested flow for conducting the communication to the above parties:

1. Preparation—Write down the Cloud and assumptions. Ensure that for A-B and AC you have strong and agreeable assumptions, for B-D and C-D' you have fortifying assumptions as well as assumptions that should be challenged (at least one of each), and that you have D-D' assumptions. Write down the amendment to the procedure you want to propose. You do not necessarily want to present your work. It may better to talk through it without showing diagrams and using TOC terminology (at least at the early stages of using TOC for managing people).

2. Meeting with the person (Shipping Clerk) who has raised the problem:

- Present the incident causing the fire.
- Present the entities of the Cloud following the sequence. A→B→D and then A→C→D'.

The sequence is based on the fact that A is commonly accepted. Then, we move to presenting B and D to radiate to the Shipping Clerk that his views are accounted and understood. Thereafter, we move to present the system view C and D'.

- Get acceptance for the logic and wording. Make notes for yourself if there are comments that should be incorporated in the Cloud.
- Ask the person for his or her ideas for permanently solving the problem.
- If the suggestion is close to your solution, acknowledge and thank the person for the contribution. If it is different, check if it is a win-win, and if it is better than the one you thought about, then you can accept it.

Otherwise, propose your solution and listen to remarks and reservations. The objective is to come up with a consolidated view for both of you.

3. Meeting with the key person (Customer Account Manager) associated with the procedure:

- Present the incident causing the fire.
- Present the entities of the Cloud following the sequence: $A \rightarrow C \rightarrow D'$ and then $A \rightarrow B \rightarrow D$.

The sequence is different as compared to the communication to the Shipping Clerk. We start with the A, and then we move to the views of the system that are represented by the Customer Account Manager C and D'. Then we move to present the views of the Shipping Clerk B and D.

- Get acceptance for the logic and wording. Make notes for yourself if there are comments that should be incorporated in the Cloud.
- Present the suggested amendment to the procedure and listen to reservations.
- If the person agrees, then you are done. Otherwise,
- If the person raises negative implications of the amendment, ask for suggestions for trimming the negative and consider incorporating them in the amendment.
- If the person comes up with an alternative idea and it is a good one, you may adopt it.
- Warning: This is one problem of many that exist in the environment. Do not allow it to become a major and

complex initiative. We need simple, practical, and rapid solutions that do not generate new problems. The check for win-win and trimming negative implications will generate good enough solutions.

We can conclude that the Fire-Fighting Cloud is probably the best tool managers can have in systematically sorting out problems that block the smooth operation of the areas under their responsibility.

Dealing with the Undesirable Effects (UDEs)—the UDE Cloud

The UDE

Before we move to the UDE Cloud, let's look at a UDE itself.

The UDE is a cornerstone in the full analytical work for developing any functional or strategic solution the TOC way.¹⁰ It is used for building the CRT that helps us to identify the core problem. Yet the UDE concept and the UDE Cloud can be used in isolation; that is, separate from the CRT. The UDE manifests a Cloud (if there is a UDE, there is a Cloud) and hence it is beneficial for the manager to reveal the UDE Cloud and use it when appropriate.

The UDE is an effect and its existence is indisputable (even though people may argue about its magnitude). It is undesirable—it endangers, reduces, or prohibits achieving a valid need, objective, or even the goal of a system.

The UDE is a cornerstone of the TOC analysis of the current reality. This is true because it focuses us on what is going wrong; that is, what it is we need to fix. It sets us on a path to changing what is undesirable to outcomes that are desirable. As such, we have to ensure that the UDE is valid

and verbalized correctly. The UDE has a clear syntax with clear guidelines:

- It is a complaint about an ongoing problem that exists in your reality and because of this problem, you cannot perform better. It should be written in present tense.
- It is a description of the state, not an action.
- It is within your area of responsibility.
- Something can be done about it.
- It must not blame someone.
- It must not be a speculated cause.
- It must not be a hidden solution to the problem (wishful thinking of solving the problem).
- It should contain one entity.
- It should not include its cause in its verbalization.
- It should be factual and not subjective.
- It should be a complete sentence.

The UDE Cloud Process

Step 1: Identify UDEs.

A problem can be defined as a UDE when:

- It has negative implications on the performance of the system.
- It has been in existence for a length of time (at least several months).
- There have been attempts to sort it out with little or no success.

Such a consistent difficulty in solving the problem indicates that the system has an inherent problem that

prevents attempts to solve this problem. We need to find what it is and for that, we need the UDE Cloud.

Another important application of the UDE Cloud is in the sales process.

Let us assume that the company has a good offer for the market based on improved service. A good offer is a solution for a problem that the potential customer is experiencing but has not managed to resolve successfully. This means that the buyer has a conflict and we had better prepare ourselves by using the Cloud method. When we are convinced that our offer breaks the customer's UDE conflict in a win-win solution, we have the basis for a value proposition for them using the Cloud.

Therefore, in this part we refer to two types of UDEs:

- System UDE—for a manager to analyze an issue within his area of responsibility
- Customer UDE—for the sales and marketing people preparing an offer to their customers

The process of building the Cloud, constructing the solution, and communicating the solution to the relevant people is identical for both types of UDEs. For the sake of clarity, the example of the system UDE is used while describing the process and the customer UDE example comes after completing Step 7 of the UDE Cloud process.

Step 2: Write the storyline.

Step 3: Build the UDE Cloud.

Building a Cloud is done through answering the questions associated with each box in the Cloud. The sequence of answering these questions for the UDE Cloud resembles a Z shape:

[B]→[D]→[C]→[D']→[A]

[B]: Why is this UDE undesirable? What important need of the system does it jeopardize or endanger?

[D]: What action should be taken to meet the jeopardized need in B?

[C]: What other important need prevents you from always taking the action D?

[D']: What action do you take to meet the need in C?

[A]: What is the common objective achieved with both B and C?

Example of a System UDE Cloud-Production

The Production Manager of an engineering company is complaining about the difficulty in assembling the final products due to shortages of parts.

UDE: We have too many shortages of parts for assembly.

Building the UDE Cloud:

[B]: What need is jeopardized?

[B]: Meet our production schedules.

If there are missing parts, we cannot assemble the product. Sometimes we start the assembly with an incomplete kit and then we have to put the unfinished assemblies to the side and wait until the missing parts arrive. In both cases, we are late in completing the products according to the production plans. This is the need of the production function as they are responsible, expected, and usually measured against the level in which they achieve the schedules.

[D]: What actions do you have to take to meet the jeopardized need in B?

[D]: Not introduce engineering changes in the schedule immediately.

Every engineering change demands the replacement of several parts. These parts are needed to be produced from scratch. It takes time to produce them. The parts for the existing design are in stock. If we wait for a while with the introduction of the new design, then we can meet our schedules.

Step 1 of the process for the UDE Cloud is different from the previous types of problems. A UDE is a well-defined problem. Step 1 is actually used to identify the UDE that we want to analyze.

[C]: What other important need prevents you from always taking the action in D?

[C]: Meet customers' requirements for the latest designs (speed to market).

Engineering changes improve the quality of our product and enhance the features they offer. In any case, the new designs provide our customers with better competitive edges. Therefore, customers really pressure us to provide them with products of the new design.

The [C] need is presented by the Marketing or Sales function. They are the custodians of the company's competitive edge and ability to sell in the market place.

[D']: What actions do you take to meet the need in C?

[D']: Introduce engineering changes in the schedules immediately.

The only way we can provide the improved products is if we introduce these new parts into the existing work orders that are planned to be assembled in the near future.

[A]: What is the common objective achieved with both B and C?

[A]: Achieve our business goals.

Our company makes money through selling products to our customers. Engineering features help the company in getting more orders in the future; effective production helps the company make money.

The system UDE Cloud is presented in [Fig. 24-7](#).

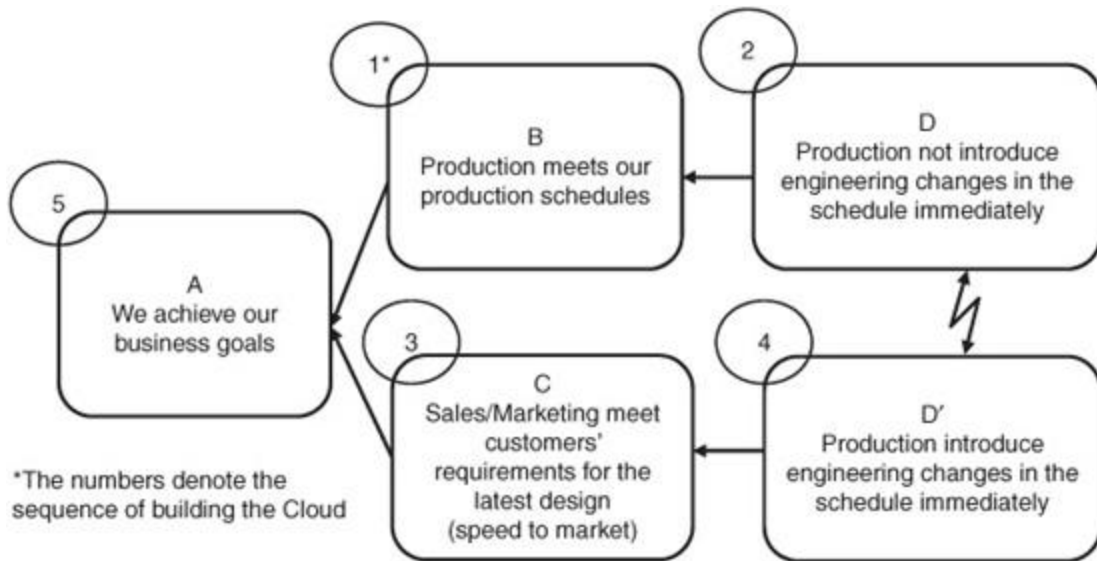


FIGURE 24-7 A system UDE Cloud example for parts shortages.

Box	Question to Guide in Writing the Content of the Box
B	What important need of the system does the UDE jeopardize or endanger?
D	What action should be taken to meet the jeopardized need in B?
C	What other important need prevents you from always taking the action D?
D'	What action do you take to meet the need in C?
A	What is the common objective achieved with both B and C?

TABLE 24-6 Sequence and the Questions for Building the UDE Cloud

In summary, the sequence and the questions for building the UDE Cloud are provided in [Table 24-6](#).

Step 4: Check and upgrade.

For the UDE Cloud, start with ensuring that the Cloud is written properly and from the point of view of the “owner” of the UDE. There is potential confusion when analyzing a customer’s UDE. Because it is *you* writing the Cloud and not the customer, the tendency is to write as a UDE the fact that the shop owner does not buy from you, does not buy enough, or is not willing to accept your offer. These are not UDEs because they imply a hidden solution: “If the customer bought from me more/more often/accepted my offer, then I would sell more.”

Another tendency is to put the new offer—the solution that you believe will improve your business—as one of the actions (D or D’). This is not the purpose of the UDE Cloud. The offer should be the injection that breaks the Cloud for the customer.

Check all logical connections of the Cloud and make the necessary corrections and upgrades.

Check that the logic of the diagonals is clear.

Tactic D is jeopardizing need C (even though this is a part of the flow of building the UDE Cloud).

Tactic D’ is jeopardizing need B.

Step 5: Surface assumptions.

Conduct the regular process of surfacing assumptions. Here are some extra points to consider.

For the system UDE Cloud, the UDE is definitely a system fault and hence we would like to surface system assumptions that may have been valid in the past, when the system was built, but may have lost their relevancy and now can cause blockage. Therefore, keep surfacing assumptions until you detect an assumption or several assumptions that

could be challenged and develop an injection to negate them.

Step 6: Construct the solution.

For the system UDE Cloud, just follow the regular guidelines for breaking a Cloud and check for win-win. The arrow, corresponding assumptions, and injections are shown in [Table 24-7](#).

Step 7: Communicate the solution.

Always prepare for the communication session with the relevant people.

In the preparation, you should consider the expected reactions and attitudes of the participants of the meeting. You should be prepared with your responses to their comments and reservations. It may be beneficial to present the subject first to at least one person who can give you feedback about the problem and the proposed solution.

Communication should follow the first three of the five layers of buy-in. You have to develop your own style and ways to handle these layers. The TP work is your homework to ensure that your views are clear to you.

Arrow	Assumption	Injection
B-D	MRP production scheduling cannot accommodate engineering changes not in the forecast.	Implement simplified DBR to reduce lead time significantly.
	Uncontrolled release of engineering changes dilutes the order.	
C-D'	The project management system in engineering is bringing the engineering changes to the market as quickly as possible.	Implement multi-project CC to reduce lead time significantly.
D-D'	The shop floor has too much WIP inventories.	Hold new product orders until all engineering changes are checked and have material requirements.

TABLE 24-7 The Arrow, Corresponding Assumptions, and Injections for the Personal Dilemma

Layer 1 achieved a clear definition of the problem through the Cloud. The consensus that in the current reality the problem outlined by the Cloud cannot be resolved should generate agreement on what the problem is.

Layer 2 is the commitment to find a win-win solution—a new set of tactics that do not conflict with each other and support the achievement of both B and C needs.

Layer 3 is the detailed injection (or injections) that breaks the Cloud. The agreement on the solution can be achieved by presenting the cause-and-effect logic that shows how the injections support the achievement of both needs.

For years, the perception was that the best way to communicate is by explicitly using the TP tools. This is not always the case. I suggest you check if presenting the TP analysis works for you and if not, find other ways to deal with the layers of consensus.

Generally, people would like to be involved in building the solution that will affect their jobs. A manager who is committed to continuous improvement first has to do the homework—to analyze and define the problem, construct a solution, and then communicate to the proper people with the view of getting their support and collaboration for adopting and implementing the solution. For communication, the manager needs to deal with the question of the improvement process, “How to Cause the Change?” The TOC subtitle for it states that we should induce the proper people to invent such solutions. “Such solutions” means solutions that are close to the solution we have found. However, they do not necessarily need to be precisely the same.

If the people come with injections that are good enough to break the Cloud, are practical, and create a win-win situation, then we should consider adopting or incorporating them, even if we have our own developed injections for the Cloud.

We will not always be able to bring the people to participate actively in proposing the solution. Our work will not always be strong enough to induce people to come forward with the logical outcome from the work that has been done. Experience will guide you in the way to communicate. Just ensure that you are flexible enough and attentive to listening to people's comments and reservations.

Example of a System UDE Cloud-Retail

Step 1: Identify UDEs.

Step 2: Write the storyline.

Step 3: Build the UDE Cloud.

Example of a Customer UDE Cloud

You are a salesperson. Your company (the supplier) is selling consumer products. For years, the company has been promoting the purchase of large quantities by offering discounts. After implementing the TOC solution of MTA (*make-to-availability*), the company adopted the mindset of "stop pushing." You want to offer your customers (shops, retailers, etc.) or potential new customers the opportunity to move to the TOC replenishment model of report daily consumption and get replenishment frequently.

You generally know the typical complaints of your customers or the customers of your competitors. You can focus on one customer in particular (that is the most relevant for your offer) and build the corresponding customer UDE Cloud. You know that your customers complain about stockouts. (Since you are building this UDE Cloud through the eyes of your customers—the shop owners—the usage of "I" or "we" in the answers to the questions implies the shop owner's position.)

UDE: We have too many stockouts.

Building the UDE Cloud:

[B]: What need is jeopardized?

[B]: Secure revenues from selling products the market wants to buy from my shop.

We [shop owners] know that we make money by selling products to consumers (the people who come to the shop). Consumers who come to our shop and do not find what they want do not generate any income for the shop. Consumers also may not come back. Hence, availability is important.

[D]: What actions do you have to take to meet the jeopardized need in B?

[D]: Buy the products that are selling well as urgent orders with special deliveries.

[C]: What other important need prevents you from always taking the action in D?

[C]: We [shop] need to control cost per unit bought.

The suppliers charge more for urgent deliveries and for smaller quantities. I [shop] can get significant discounts for buying in large quantities.

[D']: What actions do you take to meet the need in C?

[D']: Buy large quantities (even if it is more than we [shop] need for a reasonable period).

[A]: What is the common objective achieved with both B and C?

[A]: Have a successful business.

The Customer UDE Cloud is presented in [Fig. 24-8](#).

Step 4: Check and upgrade.

Let's do the check of the diagonal from the example of the Customer UDE Cloud.

Is [D] jeopardizing [C]?

In D, the shop owner wants to buy products by using urgent orders. The supplier charges more for such orders and hence the price per unit bought will go up, jeopardizing the need C.

Is [D'] jeopardizing [B]?

D'—buying in large quantities—consumes the cash reserves of the shop owner. Shops hope eventually to sell everything they have bought. However, it is only in the process of sales that it becomes obvious which products sell well and which do not. In addition, the large quantities take long to be sold. Products that are just sitting in stock do not generate money, thus endangering securing the revenues. Moreover, not enough reserves to buy replenishment of products that clients want to buy results in loss of potential revenue, which hurts B even further.

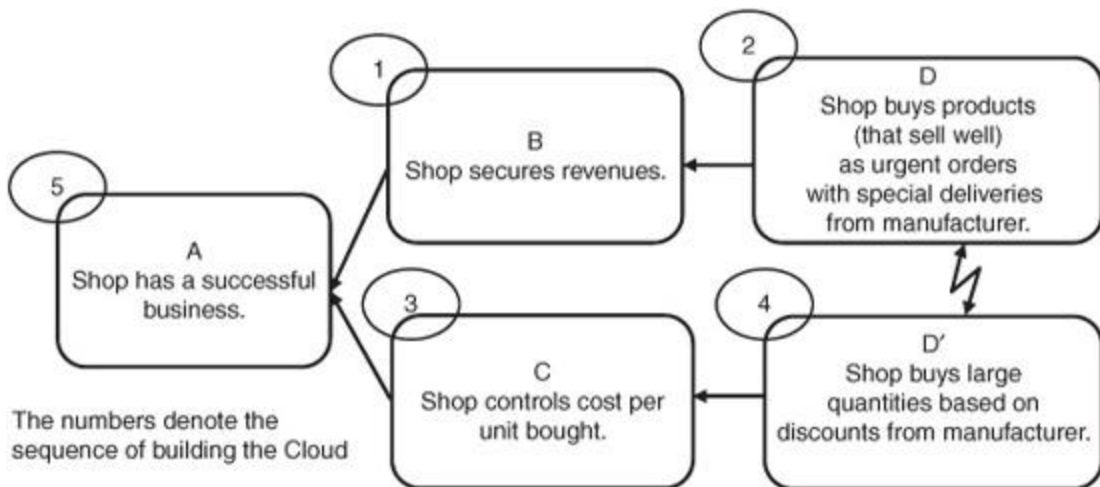


FIGURE 24-8 An example of a “Customer UDE Cloud” (from the perspective of the shop owner and we are the manufacturer).

Step 5: Surface assumptions.

For the Customer UDE Cloud, ensure that you will surface enough supporting assumptions on the B-D arrow. That will

help to secure an offer that really brings value to the customer (the shop owner). On the C-D' arrow, highlight the assumptions that the customer (the shop owner) feels reflect the policies that suppliers (you) use when determining the terms and conditions for supplying, such as minimum order quantities, shipping costs, frequency considerations, etc.

Examining and challenging these policies can provide opportunities for you to give a Mafia Offer¹¹ to your customer. The assumptions and injections are provided in [Table 24-8](#). Pay particular attention to the C-D' assumptions.

Step 6: Construct the solution.

For the Customer UDE Cloud, we recommend focusing the effort on breaking C-D'. There are two major reasons:

1. It will be easier for the customers to accept the supplier's offer that will give them what they want—urgent orders (“urgent” meaning whenever they need them and with short delivery times) without paying extra and without demanding a major change or effort from their side.
2. It will be difficult for the competitors to copy because the assumptions under C-D' reflect the common policies and business practices of the whole industry. Anything to do with mindset, policies, and procedures demands a strong determination of management and a supportive culture. It may take the competitors a long time to observe your offer, recognize its competitive edge, and agree internally on what needs to be done to catch up with you. This provides a window of opportunity for the company that pioneers offers that break Clouds on the supply side.

Let's continue our examination of the Customer UDE Cloud.

Arrow	Assumption	Injection
B-D	Manufacturers don't hold a large variety of finished goods inventory to replenish small orders quickly.	Manufacturer holds buffers sufficient to cover demand during lead time for their full complement of products.
C-D'	Manufacturers (supplier) see producing large quantities as cheaper than producing small quantities, and charge accordingly.	Manufacturers recognize that replenishing stock buffers of a central warehouse levels production, eliminates chaos, pulls raw materials evenly, etc.
	Manufacturers (supplier) prefer large quantities because of packaging considerations.	Manufacturers implement mixed shipment packaging options.
	Large batches save the manufacturers (supplier) costs of order processing (labor and computer lines).	Many pull ordering systems are easily automated, thus reducing labor and computer lines.
	Large batches save on transportation.	Increased Throughput easily offsets increases in transportation costs.
D-D'	Manufacturer cannot design a system to both respond to urgent small orders and respond to large orders at the same time.	Manufacturer implements a distribution/replenishment system to ship variety and replenish rapidly. The focus is on Throughput, not cost savings.

TABLE 24-8 Assumptions and Injections for Selected Arrows

Once C-D' assumptions (customer shop's assumptions about the manufacturer's behavior as a supplier) are verbalized, it gives the manufacturer an excellent opportunity to develop the solution that will challenge and negate these assumptions. Given that the manufacturer has implemented MTA (as per the storyline), the manufacturer can break the connection between C and D' and offer the shop *whatever quantities they want to buy, whenever the need is, and at a reasonable price*—which will satisfy both the shop's needs, [B] secure revenues through preventing lost sales and overstock, and [C] control cost per unit through making it reasonable to order small quantities.

Step 7: Communicate the solution.

For the Customer UDE Cloud, my recommendation is to develop a presentation that takes the customer through the above steps covering the problem, the direction of the solution, and the proposal based on the injection.

Addressing Multiple Problems—the Consolidated Cloud

Once we master the management tool dealing with individual UDEs, it is only natural that the manager would like to find a solution to a multi problem situation. Managers do not always have the time to conduct a complete TOC TP analysis (CRT and FRT) to develop a comprehensive solution for their area, and hence can use the Consolidated Cloud approach for a shortcut that helps construct a good enough solution that will produce short-term benefits while supporting future improvements.

We cannot call this approach a daily managerial tool, but its foundation is the one-off usages of the UDE Cloud. After using the UDE Cloud several times addressing different UDEs, you may observe a common pattern between the Clouds and you may wonder if there is a more common underlying Cloud and whether all of the UDE Clouds are derivatives of this Cloud. Therefore, you may decide one day to do a deeper analysis.

This process is also known as the *Three-Cloud approach*. Please note that the Consolidated Cloud (Generic Cloud) represents the reality around the three UDEs that have been chosen for the analysis. It is not necessarily the core problem, as the UDEs may be concentrated only in one part of the current reality and other parts may not be represented in the analysis.

We use three Clouds, as it is usually a good number to get different aspects of the subject or the area under investigation. You may decide to take more UDEs and consolidate more than three Clouds. It may help in achieving group consensus in which its members want to contribute their views on the burning problems that need to be sorted out. The general process for the Consolidated Cloud is given in [Fig. 24-9](#).

When should we use the Consolidated Cloud approach?

1. For analyzing the area under your responsibility. This is the most common use of the method when UDEs deal with the performance of the area and the behavior of the people.
2. Accelerating initiatives. Every organization has improvement initiatives. These are small projects that have been launched with the view that when completed they will bring benefit to the organization. If you are in charge of such an initiative and you are unhappy with the progress, you may consider using this approach. Just collect several of the problems that the initiative encountered and use this method.
3. BM for a POOGI. BM is a kind of problem identifier. It highlights issues that cause penetration in the buffers and cause disruption to the smooth flow of the system. The reasons for buffer penetration are collected and analyzed. We can select three typical problems, build their Clouds, and consolidate them into one Cloud.

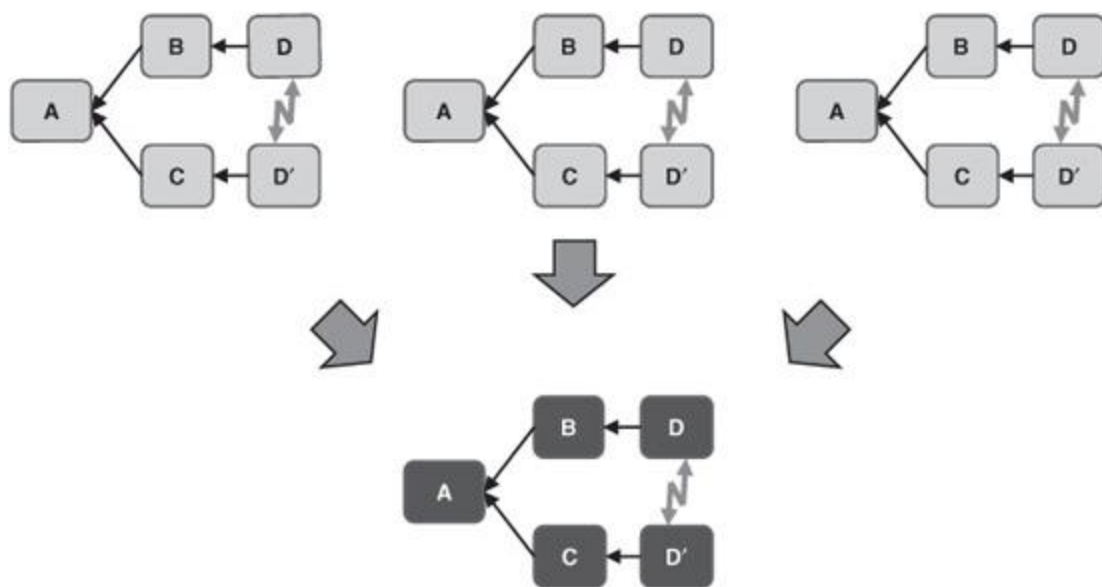


FIGURE 24-9 The general process of the Consolidated Cloud.

All of these applications can produce a Consolidated Cloud. Once we construct the Consolidated Cloud, we use it to develop the direction of the solution, a template for the injections, and specific injections to resolve the individual problems. This is a multi-injection solution for a multi-problem situation.¹²

The Process of Consolidating Process Outline

1. Select three UDEs from the area under investigation.
2. Build the individual UDE Cloud for each UDE using Steps 2 through 5 (write story lines, construct Cloud and checks, and surface assumptions) of the UDE Cloud process.
3. Consolidate the three Clouds into one Cloud.
4. Check and upgrade the Consolidated Cloud.
5. Surface the assumptions underlying the Consolidated Cloud.
6. Construct the solution and check it for win-win.
7. Communicate the solution.

Step 1: Select three UDEs from the area under investigation.

Example: A list of UDEs of a Production Manager in a *make-to-order* (MTO) environment.

UDE #1—We often do not have sufficient capacity to meet all demands.

UDE #2—Production priorities change too frequently.

UDE #3—We have too many engineering changes.

Step 2: Build the individual UDE Clouds.

In building the Clouds, recall you (the Production Manager, in this case) are always on the CD' side and the Clouds are always written from your perspective (the Production Manager, in this case).

For each UDE, build a Cloud and surface the assumptions following Steps 2 through 5 (write story line, construct Cloud and checks, and surface assumptions) of the UDE Cloud process. These Clouds are shown in Fig, 24-10a, b, c.

Step 3: Consolidate the three Clouds.

Write a generic statement in each box A, B, C, D, and D'.

- Write down each statement from the same box of each of your three Clouds. You may organize them in a small table: A statements, B statements, etc.
- Examine the statements from the same box (A, B, C, D, and D') and write a generic statement that describes all of them. Each specific statement from the same box should be an example/manifestation of the generic statement that you verbalize.

Example:

Consolidating B:

B-1: Meet our production schedules.

B-2: Effective use of resources.

B-3: Meet our cost targets.

Generic B: Meet our department performance measurements (on time and within budget).

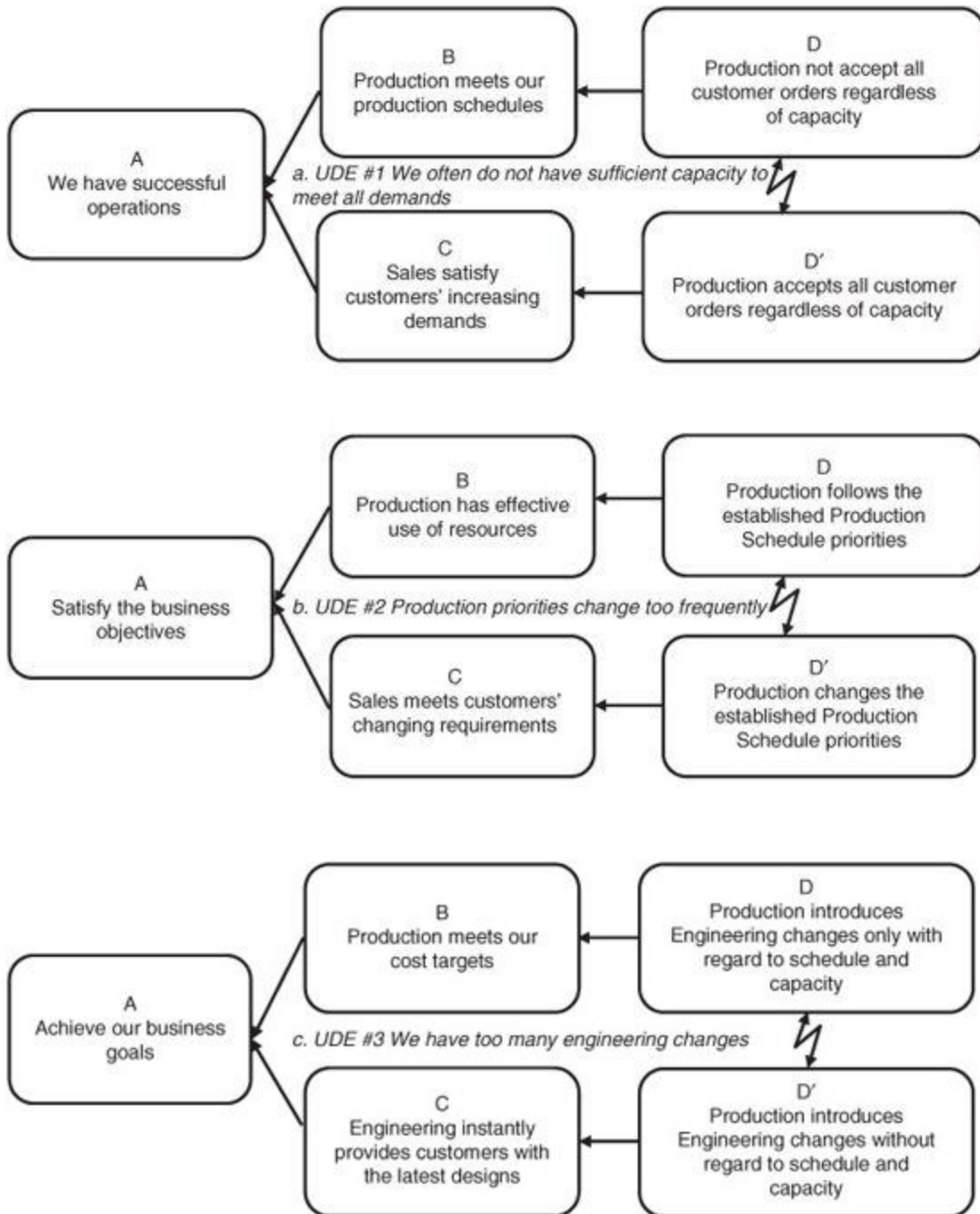


FIGURE 24-10 Examples of UDE Clouds of the production manager.

Consolidating D:

D-1: Not accept all customer orders without considering capacity.

D-2: Follow the established production schedule priorities.

D-3: Introduce engineering changes only with regard to schedule and capacity.

Generic D: Not accommodate all customer demands for schedule changes and new product introduction.

Consolidating C:

C-1: Satisfy customers' increasing demands.

C-2: Meet customers' changing requirements.

C-3: Instantly provide customers with the latest designs.

Generic C: Provide customers with flexible, fast, reliable service with the latest designs.
Consolidating D':

D'-1: Accept all customer orders regardless of capacity.

D'-2: Change the established production schedule priorities.

D'-3: Introduce engineering changes without regard to schedule and capacity.

Generic D': Accommodate all customer demands for schedule changes and new product introduction.

Consolidating A:

A-1: Have successful operations.

A-2: Satisfy the business objectives.

A-3: Achieve our business goals.

Generic A: Achieve our business objectives.

The Consolidated Cloud is shown in [Fig. 24-11](#).

Flipping Clouds

In the process of consolidating the statements of each box, you may feel as if one of the three Clouds is “flipped.” In other words, as if the B-D statements and C-D’ statements from this Cloud should swap their places to “match” the pattern we observe in the other two Clouds. If this has happened, then, for consolidation, just “flip back” these B-D sides and C-D’ sides to add them to their matching group of statements.

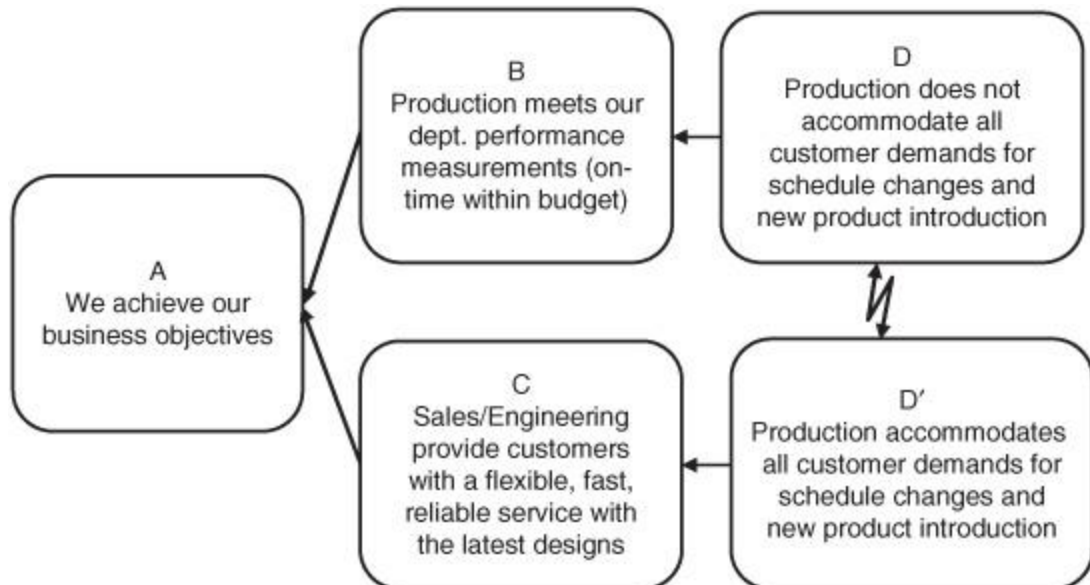


FIGURE 24-11 The Consolidated Cloud of the production manager.

Why does “flipping” happen?

The UDE Cloud is written from the point of view of the “owner” of the Cloud—the one who writes it. Thus, B reflects the need that is endangered for this person (or function).

Very often, the need in C—as recorded by this person—will be representing the need or views of another function of the organization.

However, if we are in a group consensus activity it may be that this “other function” is a member of the group doing the consolidation. He or she may see and agree with the same UDE, but from their point of view it endangers their B. Therefore, we can have the same need that appears in one Cloud in the C box and in another Cloud in the B box. Both needs have connections to their corresponding tactics. Hence, we have a situation that B-D is of the same pattern as the C-D’ of another Cloud. Knowing that this may happen, we have to review the three Clouds before starting the consolidation step.

Example:

UDE #1 We often do not have sufficient capacity to meet all demands.

In Cloud #1, this UDE was perceived as endangering need: “Meet our production schedules” that was recorded in box B.

This is a valid need of the Production Manager who is measured against meeting the production schedules.

However, from the point of view of the Sales Manager, the same UDE may endanger a different need

(that is currently recorded in box C in UDE Cloud #1):
“Satisfy customers, increasing demands.”

And thus while building the UDE#1 Cloud the Sales Manager may put in B the current wording from C, and in D the current wording from D’ and thus his Cloud will look “flipped.”

For the Production Manager, the need that is endangered is the production schedule and the other need to be considered is to satisfy the customers’ demands. The Production Manager’s view is shown in [Fig. 24-12](#)

However, when this Cloud is written from the point of view of the Sales Manager, the same UDE “We often do not have sufficient capacity to meet all demands” endangers the need to satisfy the customers and hence it will appear on the Sales Manager’s B-D side of the Cloud. The Sales Manager’s view of the endangered need is shown in [Fig. 24-13](#).

In order to consolidate the views of both managers, we have to “flip back” the sides of the flipped Cloud.

While observing the Cloud before consolidating, we can identify the nature of the needs in the Cloud. One is dealing with the needs of production and the other with the needs of the customers and presented by the sales function.

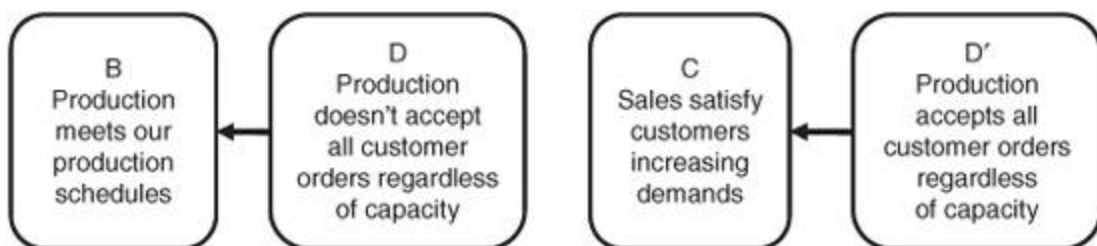


FIGURE 24-12 Example—UDE #1 Cloud from the production manager’s point of view.

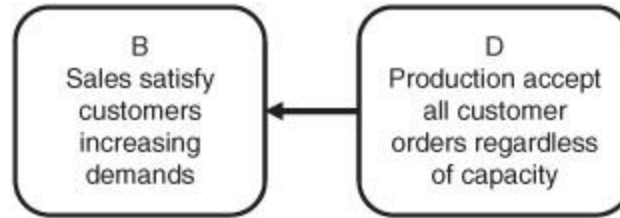


FIGURE 24-13 Example—the endangered need from the sales manager’s point of view.

The Relationships between the Consolidated Cloud and the Core Cloud

The Consolidated Cloud explains the existence of three (or sometimes more) of the chosen UDEs in an area. The role of the Core Conflict¹³ Cloud is to explain the existence of the majority of the UDEs and the inherent conflict that prevents sorting them out.

Although the Consolidated Cloud points us in the direction of the core Cloud, the analytical work that was done to reach the Consolidated Cloud may not be enough to guarantee that it is the core Cloud because the outcome of the consolidation process may be skewed by the selection of the UDEs.

The following process can be used to verify that the Consolidated Cloud can serve as a Core Cloud:

1. Take another UDE, develop the UDE Cloud for it, and check if the Cloud fits the pattern of the Consolidated Cloud. A fit means that A, B, and C are about the same verbalization and D and D’ are of the same nature of the D and D’ of Consolidated Cloud.
2. Repeat the same step for all the other UDEs.
3. If a fit is found, the Consolidated Cloud can be used as a core Cloud (if at least 70 percent of the UDEs are

represented by the Core Cloud).

4. If in the previous steps a UDE Cloud (or several UDE Clouds) does not fit the Consolidated Cloud, then a further consolidation is done by repeating the consolidation process for the UDE Clouds that do not fit together with the Consolidated Cloud. The result of this step can be called “double Consolidated Cloud.”

5. The “Double Consolidated” Cloud can be used as the Core Cloud as it represents the majority of the UDE’s of the system.¹⁴

Now that we have built the Consolidated Cloud, we move on with the process.

Step 4: Check and upgrade the Consolidated Cloud.

Step 5: Surface the assumptions underlying the Consolidated Cloud.

To be done the same way it is done for every type of Cloud.

Step 6: Construct the solution and check it for win-win.

It is unlikely that one injection can solve multiple problems and UDEs. The solution is developed in two tiers:

Cloud*	Sequence of Building	Sequence of Communicating—Always Start with A	Best Arrow to Break
Inner Dilemma	D-D'-C-B-A	A-C-D' A-B-D	C-D'/D-D'
Day-to-Day Conflict	D-D'-C-B-A	A-B-D A-C-D'	C-D'/D-D'
Fire-Fighting	B-D-D'-C-A	A -> the other side need and tactic -> the rest of the Cloud (need and tactic)	Ideally D-D'
UDE	B-D-C-D'-A	As in Fire Fighting	Ideally D-D'
Generic Cloud	A-B-C-D-D'	Start with A and then the side of the Cloud of the function representing the system that is likely to be defensive	Ideally D-D'

*Usually your side (or favorite side) is the C-D'.

TABLE 24-9 Summary of the Key Points for Each Cloud

1. Breaking the Consolidated Cloud—the chosen injection provides the direction for the solution as it deals with the general problem. This injection usually provides the necessary mindset for the solution.
2. Breaking the individual Clouds—identify injections that solve the specific UDEs to remove the specific causes for the UDEs. Therefore, we may find that one injection is not enough to solve all the UDEs and the solution will contain several injections.

Step 7: Communicate the solution.

Follow the communication guidelines as described before.

With the Consolidated Cloud procedures, we have covered the popular usages of the Cloud as a stand-alone thinking process.

Summary

Thus far, we have described five types of Clouds—the first three are for daily use to deal with one-off problems, the UDE Cloud is used for nagging problems that do not go away, and the generic (consolidated) Cloud is used for finding and addressing deeper problems and is used periodically as needed, especially for a POOGI.

A summary view covering the suggested sequence of building the Cloud, sequence of communicating it, and the recommended arrow to break is provided in [Table 24-9](#).

In the next section we will review the processes that we have used as reflected by the overall TOC methodology for problem solving—the U-Shape.

From a Problem to the Solution Implementation

The TOC process (Goldratt, 1990, 20) for identifying the problem to implementing the problem solution generally centers on responding to the following three questions:

1. What to Change?
2. To What to Change to?
3. How to Cause the Change?

The following approach is an alternative and works well on using a single Cloud to frame and solve a problem or on much larger system problems.

The TOC Methodology for Problem Solving—the U-Shape

We have covered the use of the Cloud method extensively. The suggested process for the solving problems is a derivative of the full TP methodology. Presenting the U-Shape will put all the elements together and demonstrate the way that all elements of the process are interconnected.

In a simple schematic way, the U-Shape records the logic of the relevant components that participate in the analysis of an existing current reality of a system under study (What to Change), the direction of the solution, the necessary elements of the detailed solution, and the expected benefits and impact on the performance of the system. It covers the majority of what is necessary in order to develop a full conceptual improvement solution that is viable and contains very little risk to the existing system. The structure is shown in [Fig. 24-14](#).

The U-Shape provides evidence of what is claimed to be the “inherent simplicity” of every system. Through the logic of cause-and-effect relationships, it allows the individual to better comprehend large amounts of data, to store the logical structure, and to be able to retrieve and use it when needed. It contains the TOC specially defined data elements of the system such as low performance measurements, system problems (the UDEs), the core problem, the direction of the solution, the elements of the solution (the injections), the potential risks (negative branches), and the expected

benefits from the solution, the desired effects leading to high performance measures.

The U-Shape connects the problem with the solution through the pivot—the conceptual shift from the current mode of managing to the TOC Way. Every TOC-based solution must use one of the conceptual entities of the pivot such as:

The concept of the constraint

The *Five Focusing Steps* (5FS) for constraint management

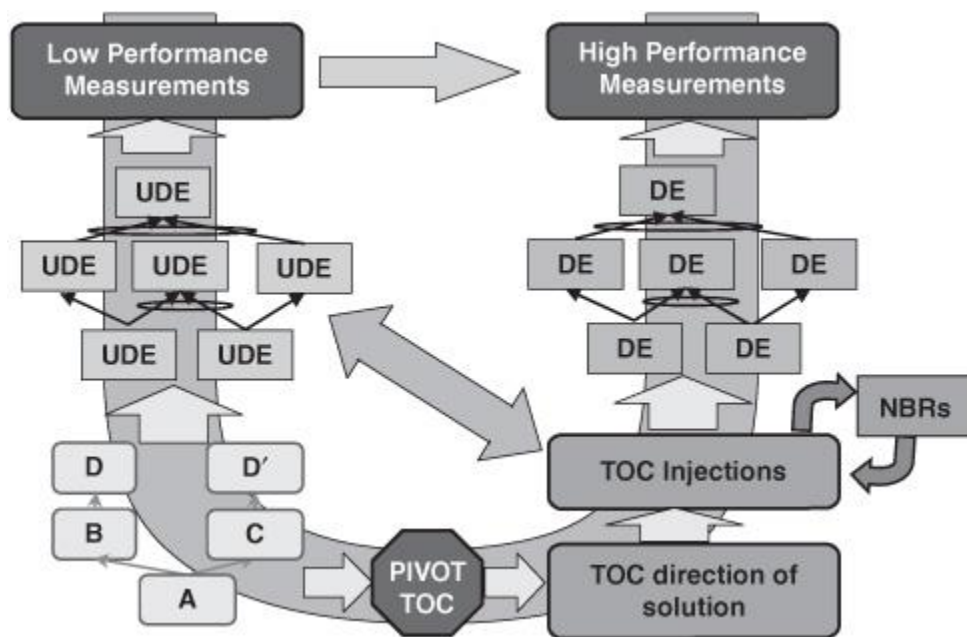


FIGURE 24-14 The detailed structure of the U-Shape.

The three basic concepts of TOC, also known as the three basic assumptions of TOC: convergence, win-win, and respect¹⁵

The process of ongoing improvement (BM, 5FS, What, to what, and how)

The U-Shape process allows the designer, the implementer, the sponsors, and the people supporting the initiative to go through a proper decision-making process that is based on a true consensus. As such, it allows the team to agree on the problem, the direction of the solution,

the elements of the solution, and their corresponding benefits.

U-Shape and the Three Basic Assumptions of TOC

The U-Shape is based on the three basic assumptions of TOC. As such, we can state what is unique about the TOC Way.

Basic Assumption 1—Convergence-reality, and specifically human based systems, are governed by cause-and-effect relationships. Hence, it is always possible to find a root cause that affects the system. The convergence is presented in the left side of the U-Shape.

Basic Assumption 2—No conflict between local and global exists. As conflicts are caused by people's perceptions or by systems, there must be a solution for every conflict. The implication of this assumption is that there should be a win-win solution for every conflict. The win-win solution is comprised of the injections on the right side of the U-Shape.

Basic Assumption 3—Handle people with respect. The entire U-Shape represents this basic assumption. It contains the respect of the managers for themselves. It reflects the seriousness in which they take their jobs. Respect for other people is demonstrated through the sharing of the work done and the willingness to ascertain and integrate inputs and views expressed by people who are relevant to the work.

The way these assumptions are incorporated on the U-Shape is given in [Fig. 24-15](#).

The overall structure of the U-Shape contains several major blocks regarding the system under study.

Current Reality—reflected in the left side of the U-Shape:

- The unsatisfactory level of performance

- The problem

Future Reality—reflected in the right side of the U-Shape:

- The solution
- Checking and removing risks
- The desired outcomes
- The improved performance

The essence of the approach for the solution point—the pivot—is the turning point from the left side to the right side of the U-Shape.

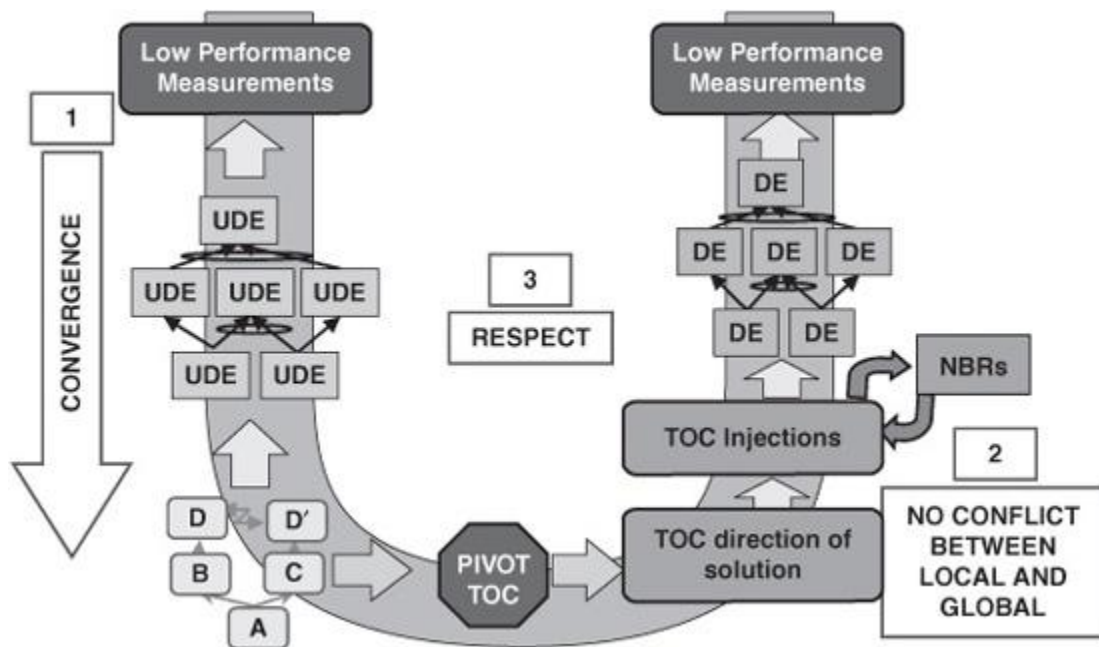


FIGURE 24-15 The U-Shape and the three basic assumptions of TOC.

Due to its generic structure of moving from the problem to the solution through the pivot, the U-Shape is valid for describing the approach for solving one problem as well as for the entire system.

The process outline used for all the daily problems corresponds to the U-Shape:

Step 1: Identify the problem—the desire to deal with the problem stems from the *unsatisfactory performance* revealed by the problem and the need to improve the situation.

Step 2: The Storyline—helps to *unleash the intuition about the current reality* explaining why the problem has caused the low performance. This is similar to *identifying the UDEs* and explaining how they cause the low performance.

Step 3 and Step 4: Build, Check, and Upgrade the Cloud— are the *manifestation of the convergence* and result in a Cloud that explains why under the current conditions it is impossible to find a workable solution.

Step 5: Surface the assumptions—is a part of working with the Cloud, in *preparation for constructing the solution*.

Steps 6: Construct the solution—corresponds to the *whole right side of the U-Shape*:

1. The pivot—in search for injections that can break the Cloud.
2. The direction of the solution—when the solution contains a change in the mindset or a system change (like in the multi-problem case).
3. The injections themselves.
4. Stating the logic that the solution will bring the benefits—the achievement of the Needs B and C of

the Cloud. These benefits are equivalent to the DEs—the desired effects. The logical connection between the satisfied needs B and C and the objective A of the Cloud leads to the improved performance of the area affected by the problem.

5. Checking and addressing potential negative effects by using the NBR process.

Step 7: Communicate the solution—The U-Shape provides an approach to communication. It captures all the knowledge that is relevant for the suggested solution. A manager who has done all the homework and has the understanding of the U-Shape can handle all comments and reservations from any person whose collaboration and support is needed. The U-Shape provides the base for a justifiable confidence of the manager in the suggested solution.

We can conclude that the process that was suggested for solving problems using the Cloud method is parallel to the methodology presented in the U-Shape. Yet there is one more element to add to constructing the solution—the NBR. This is covered in the next section.

Strengthening the Solutions—Dealing with NBRs

We have used the Cloud method to analyze the problem and to come up with a breakthrough solution—the injection. The solution that we constructed for the problem is checked for win-win, which means that we understand and can communicate the logic that supports the claim that the injection will bring the expected results and a higher performance of the system. Now that we have a potentially good solution, we should take another step—checking, addressing, and removing negative ramifications that may arise after the solution is successfully in place.

When presenting the solution to people who are closely involved with the issue, we may be confronted with Layer 4 of buy-in which stems from the fear that this good injection will also have negative side effects. This is called the NBR. As the logic of the solution is presented as a *Future Reality Tree* (FRT), the potential negative outcome is called a “branch” as if it was a “bad” branch that is growing to the side of the tree and destroys the nice shape of the good solution.

For daily problem solving, the *Negative Branch Reservation* (NBR) is used for:

1. Strengthening an injection to a Cloud—when you develop the solution and you feel that it may have some negative outcomes in the future.
2. Preparing and handling perceived negative side effects of an injection—one or more people who are directly involved with the problem and the solution may feel that the suggested injection can have a negative effect on them or their ability to perform their jobs.

Dealing with a Half-Baked Solution

When someone reporting to you suggests an improvement idea that you recognize is a half-baked solution, how do you deal with his suggestion?¹⁶ You can't say yes as it is not that good, but you can't say no as you do not want to offend a person who wants to contribute and participate in the process of continuous improvement.

The Process of Handling the Negative Branch

Step 1: Write the injection and its identified possible negative outcome in the form of a logical diagram. Place the

injection at the bottom of the page and the negative outcome at the top and check the logic by reading up from the injection: “If [injection] then [negative outcome].”

Step 2: Surface the logical arguments supporting your claim why the negative outcome is likely to happen by stating, “If [injection] then [negative outcome] **BECAUSE**”

Write down what will follow after *BECAUSE* as separate entities and decide if each new entity is something that exists now in your reality or something that will exist in the future as a result of this injection.

Step 3: If the new entity states something that will happen as a result of the injection, place this entity in between the injection at the bottom and its perceived negative outcome at the top. Now you are developing the “backbone” or “spine” of the branch. If the new entity states something that exists in your environment already, move straight to Step 4.

Step 4: If the new entity states something that currently exists in your environment, place this entity to one side of the diagram, as this will be one of the assumptions that helps explain the logic of the intermediate entity or the perceived negative effect.

Step 5: Check on the “backbone” from bottom up where the positive injection turns into a possible negative effect.

The structure of the NBR is provided in [Fig. 24-16a](#).

Step 6: Develop a supporting injection to trim the negative outcome and insert it into the diagram.

Step 7: Check that the supporting injection removes the negative outcome.

The outcome of the NBR process is shown in [Fig. 24-16b](#).

Example: Continuation of the fire-fighting story discussed in the section on Clouds.

The Customer Service Manager came with an amendment to the procedure that states that whenever shipment on time is at risk from inadequate delivery location information and the Customer Account Manager is not available, then Shipping Clerk has the authority to contact the customer about this information.

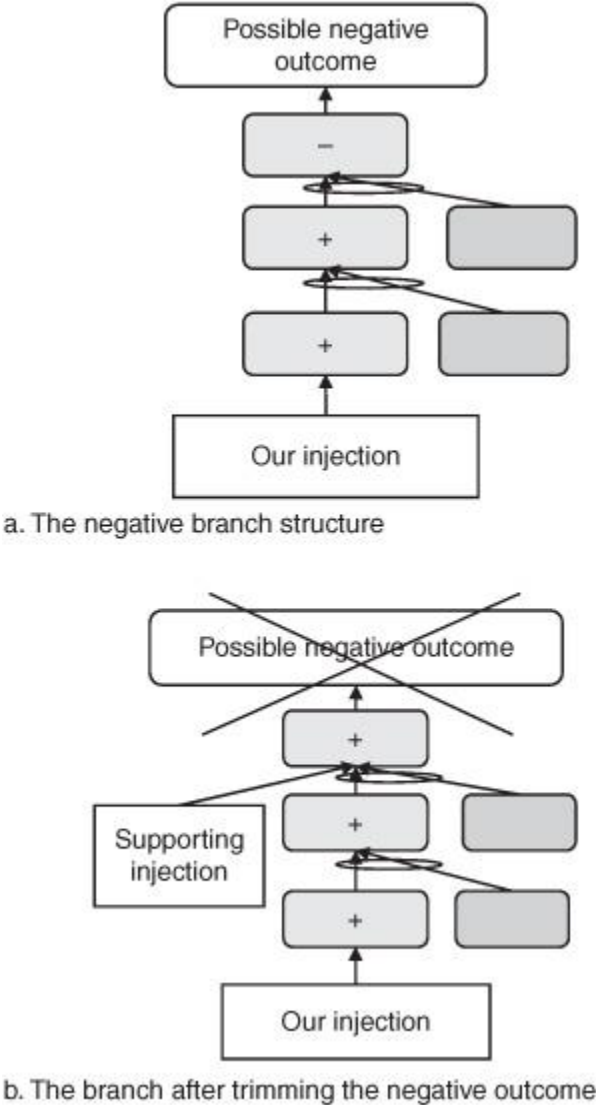


FIGURE 24-16 The negative branch and solution structure.

The manager presented the problem and the proposed injection to his team. The Customer Account Manager who was involved in this incident raised his reservation.

“Yes, but . . . if we adopt this amendment, the customer will perceive me (the Customer Account Manager) as irresponsible and unprofessional.”

Step 2: Surface logical arguments for the possible negative.

If [Injection] then [The customer will perceive the Customer Account Manager as irresponsible and unprofessional] because . . .

1. The customer will think that the Customer Account Manager did not pass on all necessary information.
2. The Shipping Clerk will tell the customer that they do not have the delivery information.
3. The customer feels that the Customer Account Manager covered all these details.

Step 3: Build the backbone with entities that will happen.

Entities [1] and [2] will happen as an outcome of the injection and hence they belong to the “backbone.” Entity [2] will cause [1] and [1] will cause the negative outcome. The logical sequence of cause-and-effect is: [Injection]→[2]→[1]→ [Negative outcome].

Step 4: Position existing entities on the side of the backbone.

Entity [3] exists in the current reality and therefore it is a supporting assumption for the causality explaining how entity [2] leads to entity [1].

It reads: IF [2] AND [3] THEN [1].

IF [the Shipping Clerk tells the customer that they do not have the delivery information] AND [The customer feels that these details were covered by the Customer Account Manager] THEN [the customer will think that the Customer Account Manager did not pass on all this information].

Step 5: Check where on the backbone it turns into negative.

The backbone turns negative in entity [1]. Entity [2] this is what is expected to be done by the Shipping Clerk when such situation arises. However, this causes [1] the customer will have the wrong perception about the Customer Account Manager and that is already negative for him.

Step 6: Develop the supporting injection to trim the negative outcome.

Supporting Injection:

The Shipping Clerk tells the customer that in order to provide the quickest, best service possible they would like to *re-check* the delivery information.

Step 7: Check that the supporting injection removes the negative outcome.

“The Shipping Clerk tells the customer that in order to provide the quickest, best service possible they would like to *re-check* the delivery information” is an action that can trim the negative outcome.

The negative branch for the Customer Account Manager with the trimming injection is provided in [Fig. 24-17](#).

However, it still may not be that great as the customer may be caught by surprise by the unexpected call and may react in an unpleasant way.

An alternative supporting injection could be:

The customer is advised in advance that on rare occasions, when the shipping instructions are not clear and the Customer Account Manager is not available (sometimes people are sick or have a personal emergency), the Shipping Clerk may call them to recheck the delivery information.

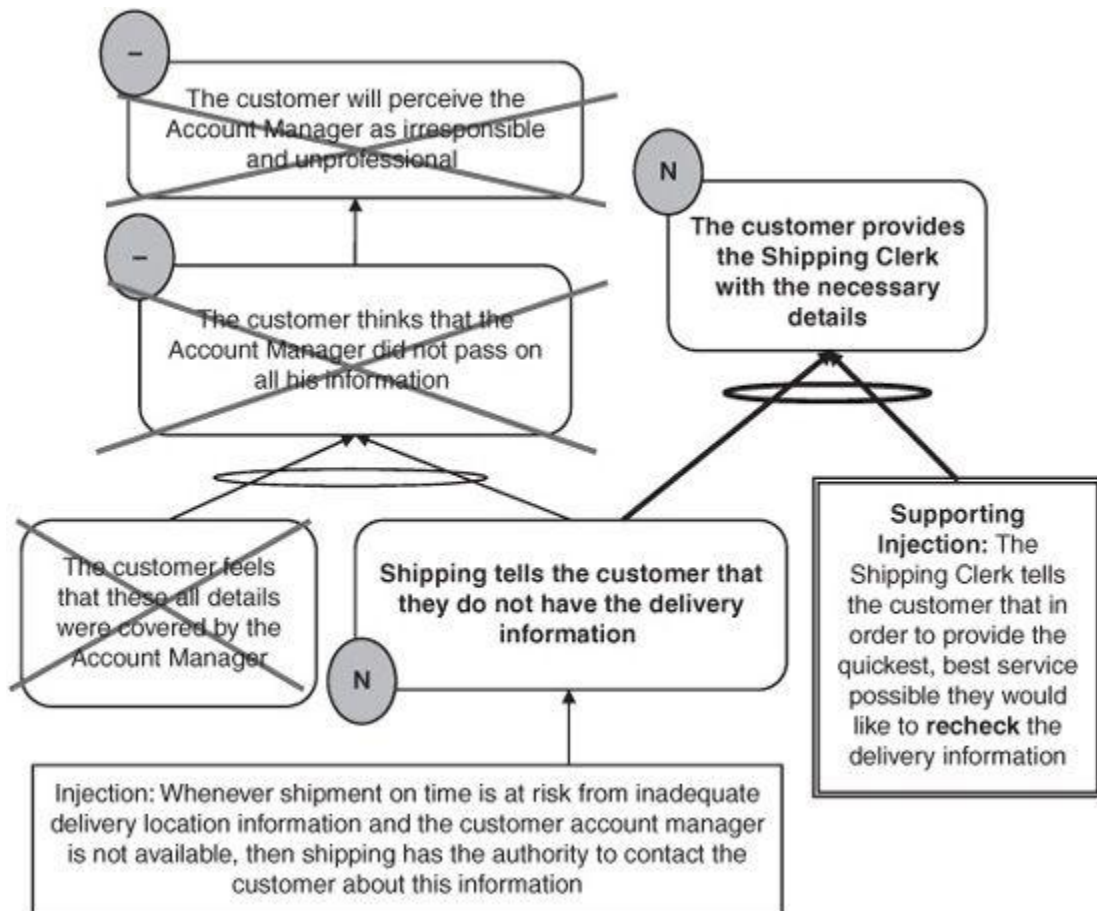


FIGURE 24-17 An example—the negative branch of the account manager with the trimmed negative outcome.

The customer may agree, disagree, or suggest ways to handle such situations. As this is discussed in advance, no harm is done and whatever is agreed with the customer becomes a part of the amended procedure.

In this section, we have seen that the NBR is another managerial tool that enhances the ability of the manager to deal with challenges—especially those that are perceived to be negative.

The issues that are raised while addressing a potential negative outcome as per Layer 4 may make the managers aware of risks that are unknown to them. On the other hand, through applying the process of dealing with NBRs, it may turn out the reservation is not substantiated and the person raising the concern may decide to drop the reservation.

The Intermediate Objective (IO) Map and Implementation Plans

Implementing an Injection—Dealing with an Ambitious Objective

The last tool of daily use of the TOC TP deals with the question of “How to cause the Change?” For daily problems that are solved by using the Cloud method, the solution is implemented mainly by communicating it to the relevant people. When dealing with fire-fighting problems, the implementation contains two stages: buy-in and the actual amendments to the procedures. When we deal with UDEs (single or multiple), the implementation also has two stages: buy-in and the change to the system or to the offering to the customers.

The implementation of an injection is an ambitious target. Therefore, we need a plan to guide us in the implementation.

There are two inputs to the planning process:

- **Necessary deliverables** in the course of the injection implementation to make sure that the injection becomes the reality. These entities are usually obtained by you or other people stating that if you want this injection to

work you must do . . . This is when experience or logic suggests clear steps for achieving such changes in reality.

- **Major obstacles** are perceived “show stoppers” that might completely block the ability to implement the injection. This input comes from the “Yes, but . . .” statements that indicate why it is going to be difficult to implement the solution in the area that we discuss. These blockages are handled with the TP that is used for building the *Prerequisite Tree* (PRT), through determining the IOs that overcome the obstacle.

These inputs are used as the building blocks of the implementation plan.

The Difference between an Obstacle and a Negative Branch

Please note that there is a difference between an obstacle that blocks our way to implement an injection and an NBR that may appear as a side result of implementing the injection. [Figure 24-18](#) illustrates the positioning of the obstacle and the NBR on the time axis of the implementation.

The Process of Addressing Obstacles

The process for addressing obstacles includes:

Step 1: Write the injection as a clear and concise statement.

Step 2: Record all perceived obstacles.

Step 3: Identify the “show stoppers.”

Step 4: Verbalize the deliverables for the obstacles that you know how to overcome.

Step 5: Develop IOs for overcoming “show stoppers.”

Step 6: Group IOs.

Step 7: Create the IO Map for implementation.

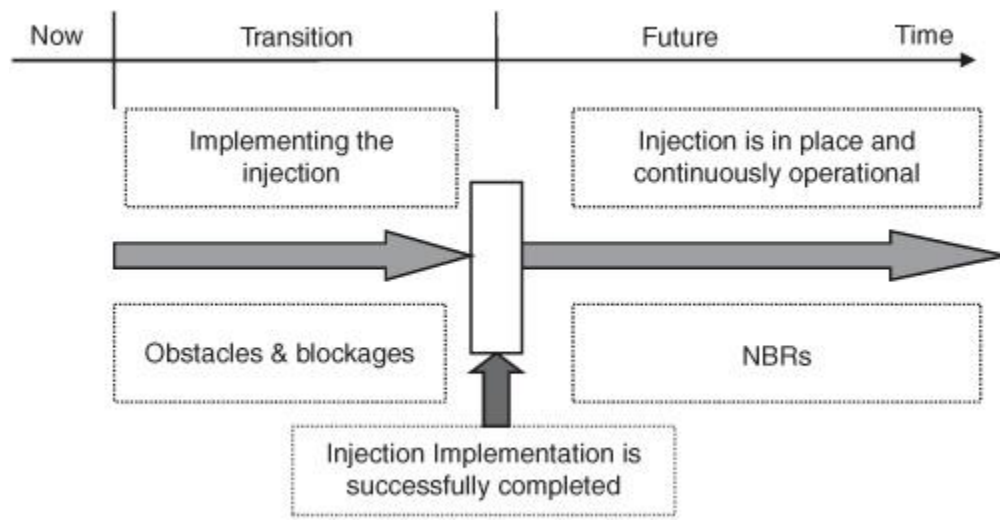


FIGURE 24-18 The relationships between NBRs and obstacles.

Step 1: Write the injection as a clear and concise statement.

Example:

An injection: A new Information System for Radiology is operational as a part of the new paperless hospital.

For Steps 2 to 5, we recommend working with a table (a Word document or an Excel file) with the following columns:

Obstacles

Show-stopper

Deliverable/IO Blocking Factor

Step 2: Record all perceived obstacles usually in the format of “we do not have” or “we do not know.”

Example:

Obstacle list (partial):

1. We do not have the scope of the implementation.
2. We do not know the acceptance criteria.

3. How do we know the system will be acceptable to all parties?
4. We do not know how to judge the quality of the converted data.
5. The existing servers are nearly full.

Step 3: Identify the “Show stoppers”.

Split the recorded obstacles into two groups (know/don't know list):

- Obstacles you know how to overcome.
- Obstacles that you do not know how to overcome—these are the “show stoppers.” Put those obstacles that you are sure that if they are not handled they will completely block the implementation of the injection in this category. You can indicate these obstacles by putting “X” in the show stopper column.

Example:

Obstacle 3 “How do we know the system will be acceptable to all parties?” is noted as a show stopper.

Step 4: For those obstacles that you *know* how to overcome, write whatever overcomes the obstacle in the format of deliverables (tangible achievements, necessary in the transition from the current situation to the full use of the injection). These tangible deliverables are in fact IOs that we need to achieve in progression to make the injection a reality. An example is provided in [Table 24-10](#).

Step 5: For those obstacles that you *do not know* how to overcome (the major obstacles or “show stoppers”), develop the IOs that you need to achieve to overcome the obstruction. Most of the time people who raise the show stoppers have ideas of how to overcome them. You need to

examine these suggestions to ensure that they help in removing the obstacles. If the IO is not that clear, you may use the intermediate steps:

1. Identify the blocking factor that causes the obstacle.
2. The major reason for an obstacle to be a show stopper is the lack of an important resource. This is the blocking factor: “We do not have money, time, manpower, willingness of our employees etc.”
3. Develop the IO to overcome the blocking factor.

No.	Obstacles	Show Stoppers	Intermediate Objectives	Blocking Factors
1	We do not have the scope of the implementation.		IO-1 We have a document that records the scope and deliverables of the system.	
2	We do not know the acceptance criteria.		IO-2 We have a document that captures the acceptance criteria as agreed with all the key stakeholders of the new system.	
3	How do we know the system will be acceptable to all parties?	X		
4	We do not know how to judge the quality of the converted data.		IO-4 We have clear indicators and procedures for converting the data.	
5	The existing servers are nearly full.		IO-5 We have a resolution for enhancing the capacity of the servers to hold all the new data.	

TABLE 24-10 Obstacles, Show Stoppers, Injections, and Blocking Factors

Example:

Obstacle 3: How do we know the system will be acceptable to all parties?

Blocking factor: Consensus on the new system (lack of consensus).

IO-3: There is a top management resolution that is based on the consensus of all stakeholders. See [Table 24-11](#).

Step 6: Review the whole list of IOs and Deliverables and if there are many, then *split them into groups related to the same topic*. An example of grouping IOs is provided in [Fig. 24-19](#).

Step 7: Sequence the IOs to create an IO Map.
Review and check the resulting implementation plan.

The IO Map

By now, we have a list of (grouped) IOs for accomplishing the injection. *The IO Map is a plan that determines the sequence of IOs to be achieved in the transition to implement the Injection.*

The logic of the sequence is that one IO has to be in place before the next IO can be achieved. There is a dependency based on tangible deliverables that each IO produces as shown in [Fig. 24-20](#).

No.	Obstacles	Show Stoppers	Intermediate Objectives (IOs)	Blocking Factors
3	How do we know the system will be acceptable to all parties?	X	IO-3: There is a top management resolution that is based on consensus of all stakeholders.	Consensus on the new system (lack of consensus)

TABLE 24-11 Overcoming the Blocking Factor

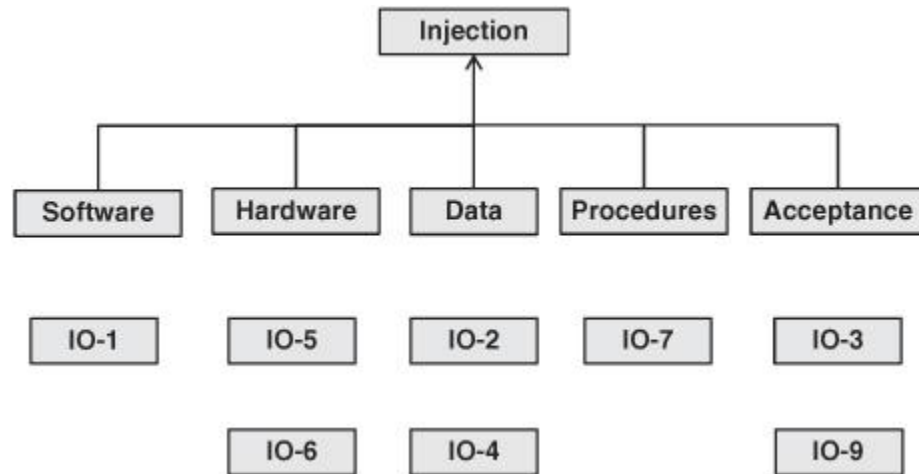


FIGURE 24-19 Example of grouping the IOs for the RIS.

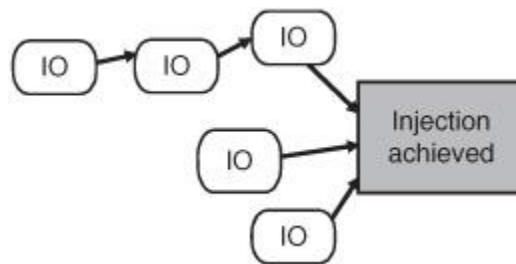


FIGURE 24-20 An example of an IO map.

The IO Map is simple and easy to construct as it is based on logic and intuition.

Sequencing the IO Map is done by stating the relationships between the IOs. Some IOs are dependent on the completion of others.

This is due to a tangible deliverable that is the outcome of one IO and is necessary for the other IO.

In the case that several actions need to be taken in order to achieve the IO, we can list them and insert them into the plan.

Here is a suggested process for sequencing the IOs:

Task 1: Copy the recorded IOs¹⁷ onto Post-Its.

Task 2: Sequence the IOs.

Start with the ambitious target on the right of the page.

Insert the IOs moving from the end (right) back to the start (left) to establish the logical dependency.

Use the method of checking the dependency between them by reading: “Before we can have (later IO) we must have completed (earlier IO).”

If there is no dependency, the IOs can be achieved in parallel.⁸

Task 3: Present the sequence to your team that has intuition about the environment. Collect the feedback and make the necessary corrections to the diagram. Check with the team that all IOs that need to be there are on the IO Map.

Task 4: Record the IO Map on an Excel file (some people tend to record the IO map on a project plan file as a PERT structure).

Implementing a Solution of Several Injections

When you have a solution that contains more than one injection, you should implement them in logical order according to the internal dependencies between them.

The sequence for creating the IO map:

1. Build the Injection Map.
2. Determine Intermediate Objectives (IOs) for each injection (as per Steps 4 or 5 of the process of addressing obstacles).

3. Sequence the IO map for each injection.

Please note: If the process is done by a group, then task 2 is done by a group check of the dependency between every two IOs as we progress for every Post-It that we introduce with an IO. The PRT is used in the full TP work for capturing the logical connection between the IOs and the obstacles as well as the reasoning for the sequencing.

4. Integration of the individual IO Maps into the Injection Map.

5. Check the integrated map to ensure that it is logically sound and complete.

Injection Map:

When the solution contains several injections, the overall plan is built by combining several IO Maps into the Injection Map of the solution.

We first build an Injection Map stating the sequence we plan to put the injections in reality. Some injections are implemented one after another; some can be implemented in parallel. Examples of an Injection Map, a fully integrated IO Map, and a Multi-Injection IO Map are provided in [Fig. 24-21](#).

The Multi-Injection IO Map can be translated into a project plan. The project plan contains deliverables and tasks.

Deliverables (IOs) are major milestones in the implementation of the injection. They are tangible and can be measured. In the plan to implement the injection, they are the intermediate objectives marking the steps toward the completion of the implementation.

Tasks are all the activities to be taken by the project team in order to achieve the deliverables. They are actions performed by specific resources and estimated time duration.

Example: Injection: *Throughput Dollar Days* (TDD) is used as the prime measurement for on-time delivery of projects. An example of a mini-project plan for implementing an injection is provided in [Fig. 24-22](#).

We can conclude that the IO Maps provide the manager with a planning tool for the implementation of the solution. Involving relevant people in the process of building the maps can create ownership and enhance the involvement and support in making the solution a reality.

Conclusion—Problem Solving the TOC Way

The TOC approach is based on the managers' self-commitment to improve the performance under their responsibility. The TOC way is to work systematically by answering the three questions of improvement (what to change, what to change to, and how to cause the change). Not every problem and challenge demands a thorough analysis and developing a breakthrough solution. Managers do make good decisions (and sometimes "pay" for their bad decisions). The purpose of this chapter was to enhance the ability to make decisions by providing tools to handle problems systematically and support in developing the skills to use them. The tools described in this chapter can be added to your personal toolbox. Practice and use them when you feel it is appropriate.

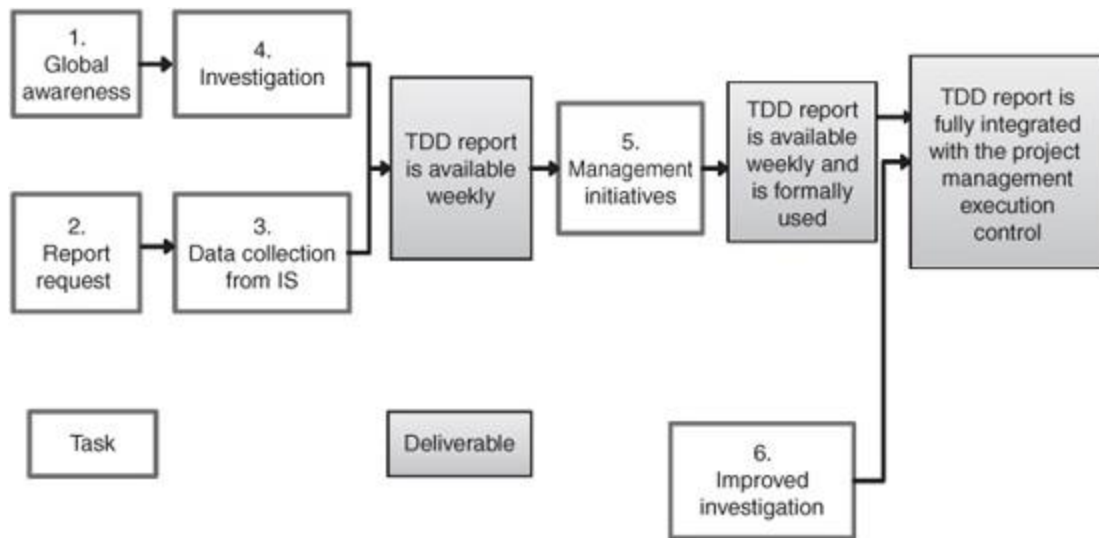


FIGURE 24-22 Example of mini-project plan for implementing an injection.

For addressing a problem systematically and explicitly, we propose the comprehensive process that we have outlined in this chapter. The flow of the process covers the three questions for improvement.

There are two inputs to the planning process:

- **Necessary deliverables** are encountered in the course of the injection implementation and must be achieved to make sure that the injection becomes the reality. These entities are usually obtained by you or other people, stating that if you want this injection to work you must do . . . Experience or logic suggests these clear steps for achieving such changes in reality.
 - **Major obstacles** are perceived “show stoppers” that might completely block the ability to implement the injection. This input comes from the “Yes, but . . .” statements that indicate why it is going to be difficult to implement the solution. These blockages

are used for building the PRT by determining IOs that overcome the obstacle.

What to Change to? Construct simple practical solutions.

- Choose an injection that breaks the Cloud and supports both needs in the Cloud (win-win).
- Deal with potential negative outcomes by using the NBR process as a part of the solution or as a part of the implementation.

How to Cause the Change? Induce the proper people to support and implement the solution (preferably by participating in the construction of the solution or by suggesting parts of the solution). To better facilitate the change, it is expected that the manager does the preparatory work (“homework”) covering the first two questions—the problem and the solution. Then facilitate the next two steps:

- Achieve consensus and buy-in.
- Develop the IO Map and implementation plan (for system changes).

If you want to be proficient with these tools, then you must continuously practice. The more you practice, the better you become and the quicker you are in using these tools, even to the extent that you can do most of the work in your head without any writing. Therefore, you must keep on practicing. Use every opportunity.

Warning: Do not push the TP on your people.

Ensure that the TP works for you but do not impose it on your subordinates. Some people may find the TP too deep, too demanding, and sometimes even threatening. Some people may feel uncomfortable with the tools

themselves and the mechanics. The TP are the tools for the individuals. I suggest that you approach your staff in stages. First, use the TP for yourself and ensure that your staff gets benefits through the way you handle and solve problems. Later on, they may be interested to know how you address the problems systematically. In the later stages, some of the staff may be interested to learn these tools for themselves. You may teach them, point them to appropriate educational materials, or send them to a school.

Embarking on TOC is a personal choice. My view of TOC is that you do TOC seriously or you do not do it at all. The strength of TOC is its knowledge and the methodology for understanding and developing new knowledge. The processes suggested in this chapter are quite demanding in terms of the amount of personal preparatory work that the TOC practitioner is expected to do. The real joy in working with TOC is making it happen. It comes when you can see that the injection is alive and kicking in the system and the people are happy to testify that the injection has brought them real benefits, proving the *Future Reality Tree* (FRT) is valid!

Improving systems performance needs a blend of three ingredients:

1. A *relevant win-win solution* is a solution that is applicable for the specific situation of the system.
2. *Leader and leadership* are to point the direction and pave the way for others to be able to move in the new direction.
3. *Supportive culture* is to provide proper subordination to the direction and to actively participate and contribute in making the vision a reality.

The design of the solution is the responsibility of the manager who adopts TOC. The other two elements are a part of the culture of the manager's area and the overall company. My suggestion is that in dealing with the improved performance of the area under your responsibility, you should adopt the approach of being firm and fair and always showing respect for people. This means that you do your homework, develop the solutions, and then communicate them to the proper people. Listen to their feedback and suggestions but do not allow the discussion to deteriorate to "analysis paralysis." You should be firm and demand closure and actions.

Last comment—I hope that this chapter has given you enough knowledge for starting your personal journey with the TOC TP. By now, you may appreciate that I have covered only a part of the vast knowledge that is there on this subject. The Cloud deserves a book dedicated to it, which I intend to write in the near future.

References

Goldratt, E. M. 1990. *What Is This Thing Called Theory of Constraints and How Should It Be Implemented?* Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. 1994. *It's Not Luck*. Great Barrington, MA: North River Press.

Goldratt, E. M. 2002. Necessary & sufficient CD-2: The basic assumptions of TOC, Goldratt Marketing Group.

Goldratt, E. M. 2009. *The Choice*. Great Barrington, MA: North River Press.

Goldratt, E. M. and Cox, J. 1984. *The Goal: Excellence in Manufacturing*. Croton-on-Hudson, NY: North River Press.

Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. TOCICO Dictionary. http://www.tocico.org/resource/resmgr/files-public/tocico_dictionary_first_edit.pdf

About the Author

Oded Cohen is one of the world's well-known names in Theory of Constraints (TOC). He has 30 years of experience in developing, teaching, and implementing TOC methodology, solutions, and implementation processes working directly with Dr. Goldratt all over the world. Among the countries to which Oded brings his expertise are the United States, Canada, Japan, India, China, the UK, Poland, Russia, Ukraine, Columbia, Chile, Peru, and many others.

Oded is an Industrial Engineer with an MSc in Operations Research from the Israeli Institute of Technology in Haifa, Israel. He was one of the developers of *Optimized Production Technology* (OPT®—a registered trademark of Scheduling Technologies Group Limited, Hounslow, UK)—the logistical software for production scheduling, the TOC Thinking Processes, and the TOC management skills.

Oded has brought his expertise to educating a whole generation of TOC practitioners and implementers. He is known for his passion for working with people who love TOC.

Since 2001, Oded has been a part of the Goldratt Group as the International Director for Goldratt Schools—the organization that is committed to ensuring that the TOC knowledge is readily available for everyone who wants to learn TOC from a teacher. Goldratt Schools plays a major role in developing and supporting TOC Application Experts and TOC Consultants who are given the knowledge and the practical know-how for implementing TOC solutions.

Oded coauthored the book *Deming & Goldratt: The Theory of Constraints and the System of Profound Knowledge—The Decalogue* and is the author of the recently published book *Ever Improve—A Guide to Managing Production the TOC Way*.

CHAPTER 25

Thinking Processes Including S&T Trees

Lisa J. Scheinkopf

Introduction: Anybody Can Be a Jonah!

If I have ever made any valuable discoveries, it has been owing more to patient attention, than to any other talent.

—Sir Isaac Newton

The *Thinking Processes* (TP) are the tools of Jonah, the beloved physicist-mentor of *The Goal's* Alex Rogo (Goldratt and Cox, 1986). In order to really gain benefit from the use of the *Theory of Constraints* (TOC) TP, you need to adapt the mentality and discipline of thinking like Jonah. You don't need to be born a genius. You don't need to have a PhD. You *do* need the conviction to think clearly, and to consider yourself a scientist. According to Dr. Eli Goldratt, "no exceptional brain power is needed to construct a new science or to expand on an existing one. What is needed is just the courage to face inconsistencies and to avoid running away from them just because 'that's the way it was always done'" (Goldratt and Cox, 1986, Introduction). This leads us to the principle on which all of TOC is based—the concept of *inherent simplicity*. Goldratt discusses this concept in *The Choice*, explaining that "*the key for thinking*

like a true scientist is the acceptance that any real life situation, no matter how complex it initially looks, is actually, once understood, embarrassingly simple” (Goldratt, 2009, 9).

Any intelligent fool can make things bigger, more complex, and more violent. It takes a touch of genius—and a lot of courage—to move in the opposite direction.

—Albert Einstein

Goldratt’s description of science and his concept of *inherent simplicity* are not new. Not surprisingly, his messages can be traced to one of the most important scientists of all time, Sir Isaac Newton. Newton’s Rules of Reasoning in Philosophy (Newton, 1729) have guided scientists since the early 1700s to recognize that “nature is simple and consonant with itself,” and thus few causes are responsible for many effects rather than the other way around; to avoid attributing more causes to an effect than are both true and sufficient to explain its existence; and to enthusiastically analyze and learn from (rather than ignore) the situations in which reality contradicts (or appears to contradict) our understanding of it (see Appendix A on the McGraw-Hill website: <http://www.mhprofessional.com/TOCHandbook>).

Copyright © 2010 by Lisa J. Scheinkopf.

When it comes to the use of the TP, people generally fall into two categories. The first consists of the people who make the decision to adapt the mentality of a scientist and the second category consists of the people who don’t. Those in the former category create meaningful improvements. They work hard at it—they exercise the muscle between their ears rigorously—but instead of feeling drained, they are energized not only by the results, but by the expansion

they have made to their knowledge and understanding of the world around them.

What are the TP tools? Why are they so effective in analyzing business and personal problems? How is the application of logic, language, and structure brought together for penetrating analysis of problems and conflicts? How do the TP tools then help in laying out the transition from an undesirable present to a desirable future? How do they help protect a plan from unanticipated pitfalls? How do they link together as an integrated system of logical capabilities for bringing about positive change? I hope to answer these questions in a way to show that almost anyone willing to do the work can achieve deep insight and make significant and meaningful improvements to environments both simple and complex; with step-by-step instructions on how to do it.

I begin with discussion of the tenets in logic and fundamental assumptions in philosophy that underlie the TOC TP. Then I illustrate how the discipline of diagramming helps in guiding our analysis. Each of the TP tools is discussed in sequence with instructions on how to use it. The chapter moves on to examples, some of them real application cases.

The Basic Building Block—Cause-and-Effect Logic

You see there is only one constant. One universal. It is the only real truth. Causality. Action, reaction. Cause and effect.

—The Merovingian, *The Matrix Reloaded*
When we accept the premise of *inherent simplicity*, we accept the premise that every element of a system is connected to the system via cause-and-effect relationships

with the other elements of the system. This means that the better our capability to uncover and understand the actual cause-and-effect relationships that exist today, or that we intend to put into place tomorrow, the better our capability to improve.

What do we mean when we say there is a cause-and-effect relationship? We mean that by the mere fact that one condition exists in a system, another condition is an inevitable result. Let's look at a simple example which may seem trivial because it is obvious, yet it does illustrate clearly the basic building block of the TP.

It is evening, and you have just arrived home from a day at work. You open the door to your home and turn the switch that operates the lamp in the hallway to the "on" position. The lamp doesn't turn on. What could be the reason? After verifying that you did in fact turn the switch to "on" rather than "off," you check to see if the lamp is plugged in. Why? Your life experience has led to your intuitive understanding of a cause-and-effect relationship—you know that if the lamp is not plugged in, the light will not turn on.¹ You find that the lamp is not plugged in. Aha! You confidently plug the cord into the wall, flick the switch on again, and—oh no, the light is still not on. What do you check next? Your brain goes through a quick checklist of potential causes for the light not turning on. Do you change the light bulbs? Do you turn on another light in your home to verify that the problem is isolated to the lamp and not a larger issue such as the circuit breaker or fuse, or even an electricity outage in the neighborhood? Any of these would be sufficient to cause the lamp to not turn on, so you keep checking—in the order that your intuition, which is based on experience with similar situations, tells you is most likely to least likely—until you uncover the cause, make the appropriate change, and turn on the light.

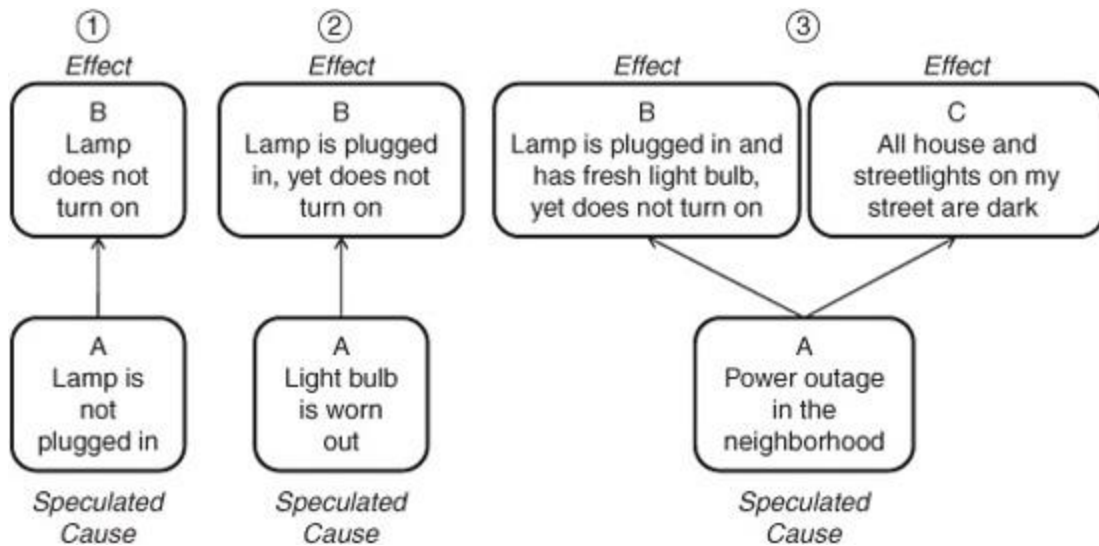


FIGURE 25-1 Cause-and-effect map—lamp does not turn on.

[Figure 25-1](#) graphically illustrates the cause-and-effect map you built in your mind. Please note that as you gained more information, your cause-and-effect mental map enlarged and you better understood the situation. You checked directly the facts you could check directly, and you modified the “entities”—your verbalization of the facts—as you went along. In the third scenario, when you finally looked outside at the rest of your street and found that it, too, was as dark as your lamp, you predicted and verified an effect that gave credence to a potential cause. If the street lights and neighbors’ lights were on, you would continue checking for alternative causes.

You also may not have been satisfied that you had at last verified the cause—you may have decided to speak with a neighbor or call the utility company. If they did in fact verify the power outage, the resulting cause-and-effect map would have looked like [Fig. 25-2](#).

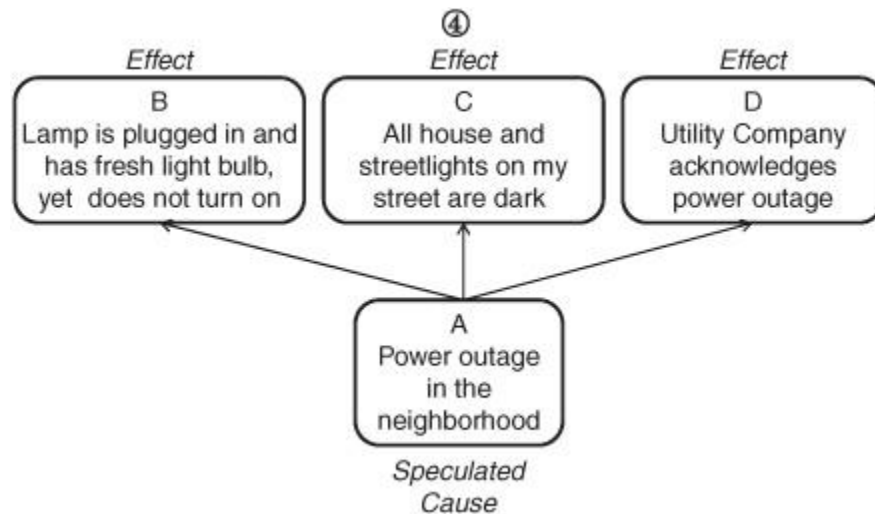


FIGURE 25-2 Cause-and-effect—power outage in the neighborhood.

In this example, you instinctively conducted checks on the hypotheses of cause-and-effect you were making, and you used a process to do so.

1. You identified a problem. *The light doesn't work.*
2. You hypothesized a cause. *The switch is not turned on.*
3. You checked your hypothesis by checking for two conditions:
 - a. You verified the condition. You checked to see if “switch is not turned on” was actually the case. It was, in fact, turned on, so you hypothesized a different cause, and then verified that the condition existed.
 - b. You validated the cause-and-effect connection. Was the fact that the lamp was not plugged in really the cause for the lamp not turning on? You checked directly by plugging in the lamp and it still did not turn on! So, back to hypothesizing a condition that could cause the lamp to be out and then validating the cause-and-effect connection.

When you adapt the mentality of the scientist, you will do these checks automatically. As we make our way through the chapter, we will expand our understanding of these a template for the detailed process of checking is provided.² It is also provided in Appendix B, which is located at the end of the chapter for your convenience.

While the example I used may seem trivial, the scientific process is not. Most of us simply are not practiced in using or communicating cause-and-effect logic. Dr. Goldratt recently conducted an experiment. He asked about 40 people—all were intelligent, educated adults ranging in age from 20-something to 60-something, ranging in professions from student to CEO—to think of and then write a sentence that contained the word “because.” The only qualifier for the sentence was that it needed to be a sentence that the individual writing it believed. In other words, they were each asked to make a statement of cause and effect that they believed to be correct. There were a wide variety of sentences, such as “I discipline my children because I care about their well being” to “Americans drive SUVs because they don’t care about the environment” to “My boss and I don’t get along because . . .” to “The cake tasted bad because the recipe was lousy.” Dr. Goldratt then asked the group to apply the simple checks to their statements. In the vast majority of cases, the individuals wrote to him and said that once they applied the checks, they came to realize that their original statements were *wrong*.

Think about how many decisions are made every day based on assumptions of cause and effect. If the group of 40 is any indicator—and I have no reason to believe they are an exception to the general population—I cannot help but think how many decisions are wrong. People are hurt and organizations do not improve, due to our carelessness in the use of “because.” The only difference between using cause-effect thinking in a situation like the lamp and a situation in

which the direction of an organization is set is the decision to really check the assumptions that would drive a given course of action.

When you develop the habit of using cause-effect, using it to make the tough decisions will be as natural as using it to figure out why the lamp does not turn on. I cannot stress the importance of practicing—of exercising your brain muscle to think clearly, and to regularly map the cause-effect statements you use, hear, and read (the sentences you use that contain the word “because”). This is the best preparation you can do for when you need to reach for the TP to make the big improvements you care about. By incorporating into your daily practice the use of the basics that I introduce in the next section, you will have everything you need to use—and even develop for yourself—the TOC TP.

Basic Terms and Mapping Protocol

Cause and effect are two sides of one fact.

—Ralph Waldo Emerson

An *entity* is the description of an element of the situation. An entity can be an effect or a cause. Keeping in mind our desire to think and communicate clearly, entities are stated as simple and complete sentences. As we make our way through the various application tools, we will identify special types of entities. Note that an entity is not a statement of cause-effect, which is a description of the cause-and-effect relationship between at least two entities.

An *arrow* is used to illustrate a cause-effect relationship between two entities. It is the graphical representation of the word “because.” The entity at the pointed end of the

arrow is the effect, and the entity at the nonpointed end of the arrow is the cause (see [Fig. 25-3](#)).

An *And Connector*³ is an ellipse or a straight line across the cause-and-effect arrows used to illustrate a “logical and” relationship between multiple entities that together form a single cause for an effect. All entities that are “captured” by the “and connector” are required as causes for the effect to occur. To better understand “logical and,” see [Fig. 25-4](#).

Entity B is an effect of both entities A and C. Neither Entity A nor Entity C can cause Entity B alone, both must exist. Moreover, when both exist, Entity B is an inevitable result. Let us use a simple example. It is your friend’s birthday, and you, along with a group of his other friends, have decided to make a surprise party to celebrate the occasion. You are all gathered in his home, and the big moment arrives. He opens the door, walks in, and you all jump up and shout, SURPRISE! Is he surprised? Yes, but only if he was not expecting the party. See [Fig. 25-5](#) for an illustration of the cause-effect involved. Note that if either of the two causal entities did not exist, he would not be surprised by any one of them.

[Figure 25-6](#) illustrates a simple cause-effect tree. There are 12 entities and 8 cause-effect relationships. Of the 12 entities, 5 are causes only, 2 are effects only, and 5 are both causes and effects. Can you identify the entities, causes, effects, and cause-effect relationships depicted in the tree?⁴

We have already established two of the fundamental assumptions of TOC: the concept of *inherent simplicity* and that anybody can think like a scientist if they choose to do so.

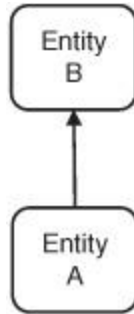


FIGURE 25-3 Entities.

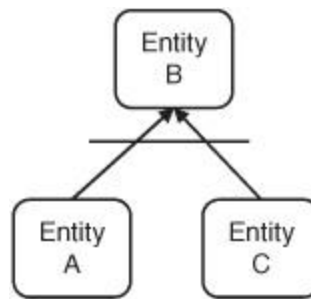


FIGURE 25-4 The “and” connector.

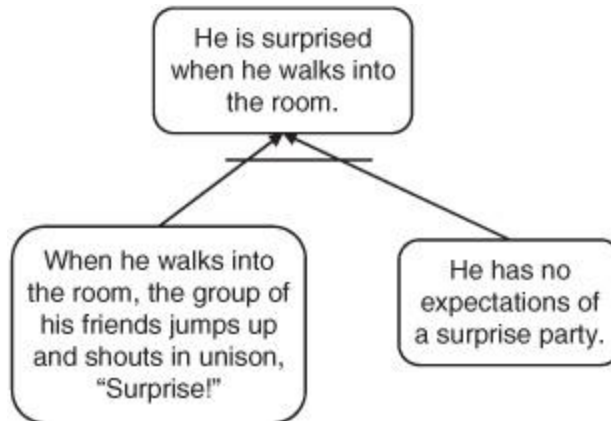


FIGURE 25-5 Example of “and” connector.

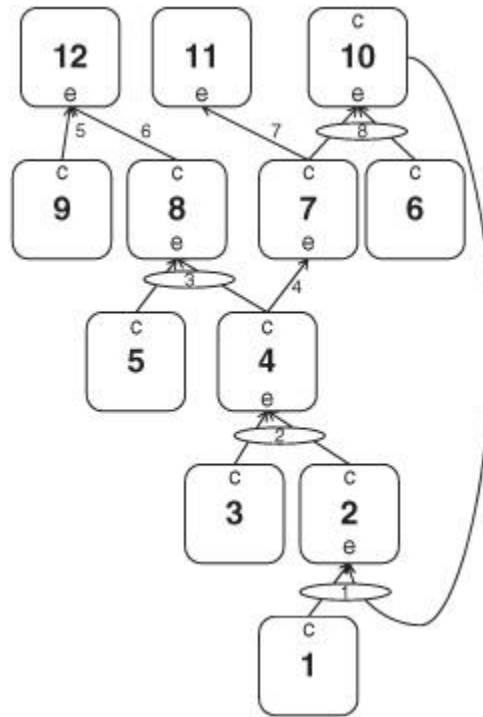


FIGURE 25-6 A simple cause-and-effect tree.

When I say “fundamental assumptions,” I mean that these are two entities that TOC takes as “facts.” With just these two assumptions as our guide, we can bring to light three more very important pieces of the foundation on which all of the powerful TOC applications are based, and on which your use of the TP will be most productive and beneficial:

1. People are good.
2. Every conflict can be removed.
3. There is always a win-win solution.

Please refer to [Fig. 25-7](#), which is a small cause-and-effect tree that illustrates how these three basic elements of TOC are derived.

Start at “the bottom” of the tree, at Entity 1, which summarizes the essence of the concept of inherent simplicity. When we couple that with Entity 2, the definition

of “conflict,” it becomes obvious that “conflict” is not a natural state, and thus must be man-made (Entity 5, given the definition of “man-made” in Entity 4). Now go to the left side of the tree. Again, we start with the summary of the concept of inherent simplicity in Entity 1. If you agree that human beings are actually part of nature (Entity 6), then it would become obvious also that our natural state as human beings is, as described in Entity 7, harmonious—consonant with the rest of nature, in harmony with ourselves and other people. It is no wonder, then, that Goldratt insists, “people are good” (Entity 8). Entity 11 states that people have the innate ability to think logically. When we combine this with what we have by now established—that people are naturally harmonious and conflicts are man-made—we have no choice but to recognize that people have the innate ability to eliminate conflicts (Entity 9) and the innate ability to create harmonious solutions (Entity 12). The result of these are the TOC premises (verbalized in Entities 10 and 13) that “every conflict can be removed” and “there is always a win-win solution.” I encourage you to study this tree, and to use it for practicing your own use of cause-and-effect logic. Would you add or modify any entities? Are the causalities solid? What tests would you conduct to verify the entities or validate the causalities represented? If you agree with the tree, what else stems from it? Can it help you to explain any of your own life experiences?

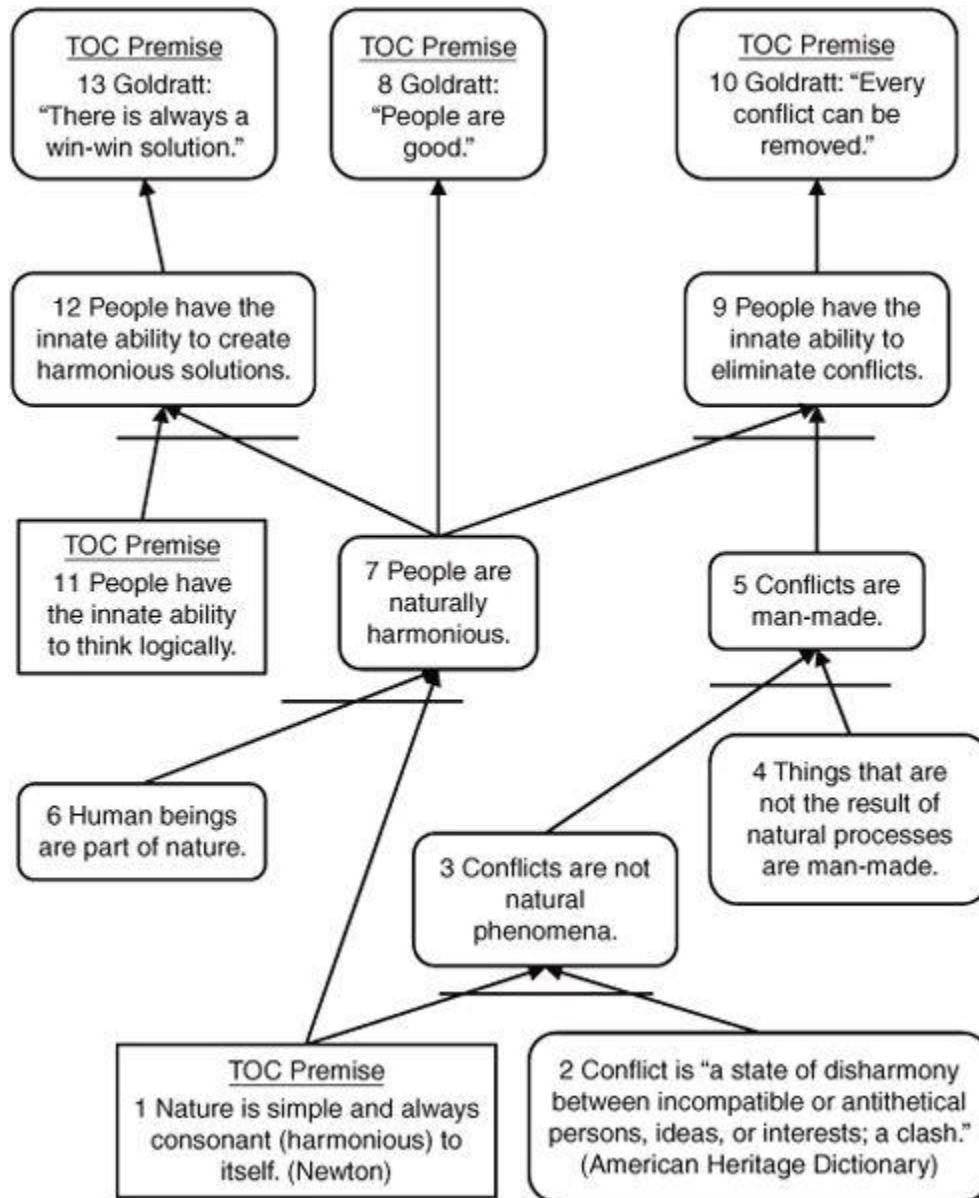


FIGURE 25-7 Deriving the three basic elements of TOC.

We are at a crucial point in your TOC TP education. We have logically derived some fundamental concepts that TOC views as “facts,” which formulate basic principles guiding the use of the TOC TP tools:

1. The concept of inherent simplicity: Nature is simple and consonant (harmonious) to itself.

2. People are good.
3. People have the innate ability to think logically.
4. Every conflict can be removed.
5. There is always a win-win solution.

I guarantee that your use of TOC will be much more fruitful if you use these five principles to guide your way. It is also likely that you are not so convinced that they are “facts.” I would ask you, then, to simply agree that they are a *possibility*. Once you agree that they are a possibility, and you consider just the possibility when you go about your daily problem solving, then I have little doubt your use of the TP will be worthwhile for you.

The last of the human freedoms: to choose one’s attitude in any given set of circumstances, to choose one’s own way.

—Viktor Frankl

The rest of the chapter is devoted to teaching you the various “standard” TOC TP. We start with tools that can be used to help you become more productive on a day-to-day basis, and then we move into the tools that are used in a “full analysis”—the systematic approach to answering the three questions of change. Please note that all of the “standard” TP are simply applications of what we have covered thus far in this chapter. If you read no further, and simply put into practice what we have covered up to this point, *you* would have the ability to derive the tools yourself when the need arises.

Tools for Daily Decision Making and Problem Solving

While we are free to choose our actions, we are not free to choose the consequences of those actions. Consequences are governed by natural law.

—Stephen Covey

Everything we do, every action we take, places a cause into reality and the effects (results) of the cause (our action) inevitably happen. The results (effects) of our actions do not have a choice, but the actions we take (the causes we put into motion) *are* a result of the choices we make. An action is putting in motion a conscious or not-so-conscious decision. Whether we are consciously or not-so-consciously doing so, we are making many decisions every day, day in and day out. Many of the decisions we make not only impact us personally, but also have an effect on others—our partners, families, teammates, associates, clients, suppliers, shareholders, communities, etc. Of course, the decisions made by others quite often have an effect on us.

Living is a constant process of deciding what we are going to do.

—Jose Ortega

Negative Branch Reservation (NBR)

We can evade reality, but we cannot evade the consequences of evading reality.

—Ayn Rand

Think about how often well-intentioned actions have led to undesirable consequences. The *Negative Branch Reservation* (NBR) is *the standard TOC TP tool with which we use cause-and-effect thinking to predict, as best we can, the effects of a given cause (e.g., action), and modify our*

idea before taking action in order to prevent undesirable consequences of taking the action. Situations in which the NBR is most commonly used are:

- Someone has presented you with an idea that they think is great, but from your vantage point, you see potential problems stemming from it. (You are thinking, “Yes, but . . .”)
- You are presenting (or preparing to present) someone with an idea you think is great, but from their vantage point, they see (or might see) potential problems stemming from it. (They are thinking, “Yes, but . . .”)
- You have an idea, and your intuition is telling you that your idea is still incomplete. (You are thinking, “Yes, but . . .”)

The NBR maps the cause-and-effect relationships between an idea (the cause) and the *undesirable effects* (UDEs) that are predicted to stem from that idea (cause). It is then used to modify (typically by expanding on) the idea in ways that would prevent the UDEs from becoming reality.

With the NBR, we introduce the entity type *injection*. An *injection is an entity that describes an element of an idea (solution) that is intended to be implemented*. Injections are always entry points to a tree such as the cause-effect trees just discussed. They represent elements of the system that do not yet exist in the system, but that will be consciously *injected* into the system in order to cause the changes desired.

[Figure 25-8](#) illustrates a simple NBR. Note that the only entry points to the tree (entities that are causes only) are either elements of the system that exist today (and therefore can be checked to exist in the system today) or injections (elements of the system that do not exist today but are intended to be injected into it in order to cause the change). Every entity that is an effect (entities that have at

least one arrow pointing into them, whether they are also causes and have arrows pointing from them) is stemming from an injection, and thus does not exist in the current environment. Therefore, these entities are *predicted* to become part of the future state of the system.

I want to stress the importance of considering the reason that you or others have generated the idea in the first place—the benefits that the idea, once implemented, are intended to produce. Acknowledging these benefits will provide you with the stamina to work through the negative branches of your own ideas to achieve the benefits. And, to communicate your reservations about other's ideas in a way that they will understand you are not trying to throw out their entire idea and its benefits, you just want to trim the potential negative ramifications. As a result, you will foster a spirit of collaboration rather than confrontation.

Constructing a negative branch is simply using the rules of cause and effect to clarify, validate, and resolve a concern over a potential negative ramification of an idea. The major steps are:

1. Write the idea as an entity. If there are multiple elements of the idea, try to write each element as a separate entity. Often, it is just one or two aspects of the idea that are responsible for the concern and this will help you illuminate only the problematic elements of the idea.

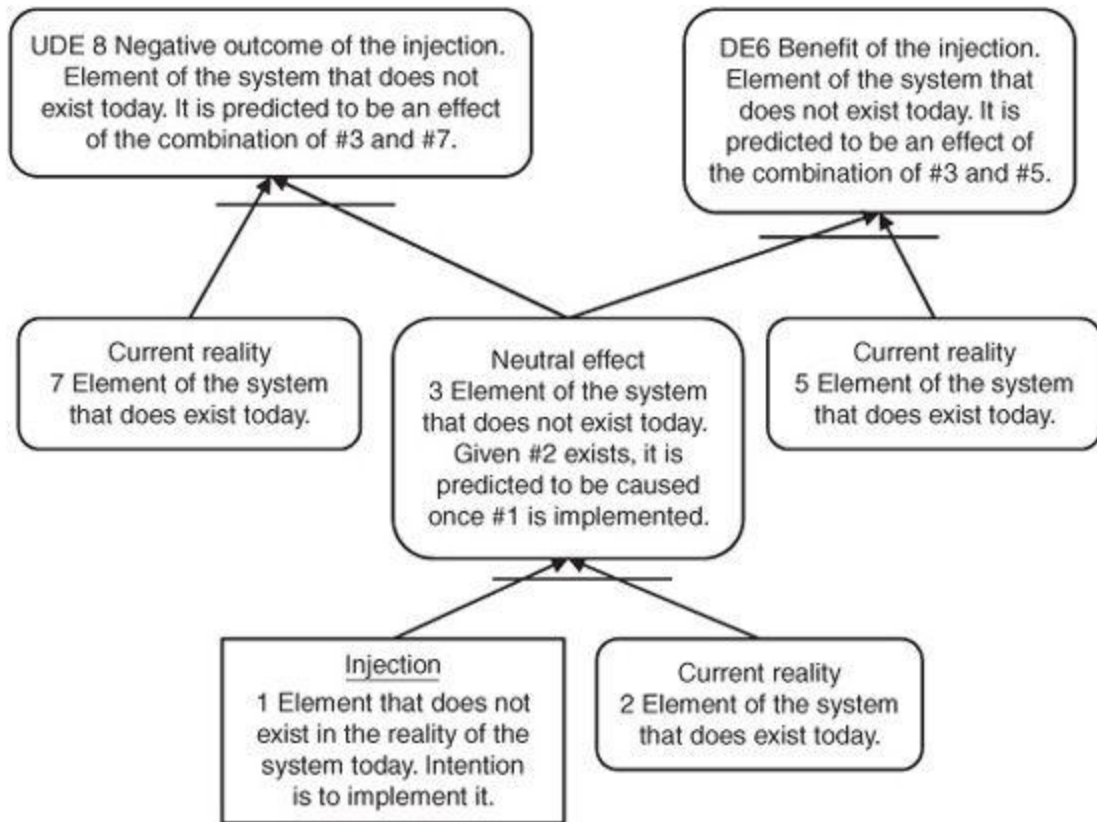


FIGURE 25-8 Simple NBR.

2. Make a list of the pros (benefits) and cons (concerns) of the idea. Write the negative outcomes that you are predicting as entities—these are the predicted UDEs.⁵ Again, try to write each element as a separate entity. Your list of cons of the idea may contain two types of concerns:
 - a. The first type of concern is consequences that would occur once the idea has been implemented. This is the type of concern that the NBR addresses.
 - b. The other type of concern is an obstacle. In this case, the concern is not with the idea itself, but rather with things that would get in the way of implementing it. The TOC TP tool that is used to deal with obstacles is the *Prerequisite Tree* (PRT), which will be described later in this chapter.⁶

3. Using the mapping protocol discussed earlier in this chapter, connect the injection entity (or entities) using cause-and-effect logic to the predicted UDEs. If you are predicting several UDEs, you may choose to build a single NBR that would encompass some or all of the predicted UDEs, or a separate NBR for each predicted UDE.
4. Check the validity of the cause-and-effect relationship and make adjustments so that it reflects your full hypothesis. This effort will likely lead you to add additional entities and layers along the way, as you make your concern clearer and clearer through the mapping process. Refer to the simple checking process discussed earlier in the chapter.
 - a. Verify the existence of the causal entity. An NBR is triggered by some aspect of the current reality that, when combined with the future that is going to be created, will hypothetically cause the undesirable consequences. What is that condition, and does it really exist?
 - b. Validate the cause-and-effect connection between the hypothesized cause and the predicted undesirable consequence. There are usually simple “mind-experiments” you can do, which would either prove the hypothesis wrong or add confidence in its validity.
 - c. Don’t be surprised if you find that a key assumption you were making was actually incorrect, and you discover that the idea would not (or most probably would not) lead to the negative outcome with which you were initially concerned.
5. Now it is time to “trim the negative branch.” Identify the place in the tree where the transition from “neutral” to “negative” occurs. In [Fig. 25-8](#), this would be where entities 7 and 3 cause entity 8. It is at this

intersection where we identify an additional idea that, if implemented, would either prevent 8 from occurring, or even replace 8 with an effect that would become an additional benefit of the solution. Check to make sure that this new, added injection does not lead to more ramifications that are negative. If it does, either replace it with a different injection or add an additional injection to trim the new negative branch.

In [Chapter 24](#), Oded Cohen provides detailed step-by-step instructions for constructing and solving negative branches. A great example of a negative branch is in [Chapter 8](#) of Eli Goldratt's book, *It's Not Luck* (1994, 53-58). I will also provide an example of an NBR later in this chapter, when I review the use of a Strategy & Tactic Tree.

Evaporating Cloud (EC)

A cloud does not know why it moves in just such a direction and at such a speed. It feels an impulsion . . . this is the place to go now. But the sky knows the reason and the patterns behind all clouds, and you will know, too, when you lift yourself high enough to see beyond horizons.

—Richard Bach, *Illusions*

The second standard TOC TP tool that is used on a regular basis is the *Evaporating Cloud* (EC).⁷ The Cloud is the tool that enables us to eliminate any conflict, and paves the way for a win-win solution. In a world where conflicts do in fact exist, and in which nearly everyone believes that the only way to deal with a conflict is to compromise (which typically means that all parties settle for less than what they really need in order to “meet in the middle”), why is TOC so bold to claim that every conflict can be eliminated?

We look no further than the concept of inherent simplicity for the answer. *A conflict is a situation in which each side thinks that it needs something that is in direct contradiction with (cannot coexist with) what the other side thinks that it needs.* If we accept Newton's statement that nature is "always consonant (harmonious) to itself," then we must accept that in reality, there are no *real* contradictions. It must be, then, that any conflict contains an erroneous assumption that blocks the ability for each "side" to get what it needs, and is thus blocking what should otherwise be a naturally harmonious reality.

Eli and Efrat Goldratt provides an excellent explanation in *The Choice* (Goldratt, 2009, 46-47).

Suppose that we have two different techniques to measure the height of a building. And when we use them to measure the height of a specific building we get two very different heights. Facing such an apparent contradiction no one would say, let's compromise; let's agree that the height of this building is the average between the two measurements.

What we would say is that somewhere along the line we have made an erroneous assumption. We'll check to see if, in the time that passed between the two measurements, additional floors were added. If that's not the case, we'll explore if our assumption—that each of the measurements was carried out properly—is correct. If they were, we'll look for an erroneous assumption in the techniques themselves; we'll explore the possibility that one of these two techniques is faulty. In extreme cases, we'll even doubt our understanding of height. But we'll always look for the erroneous assumption and never contemplate the possibility of compromise. This is how strong our belief is that there are no contradictions in nature.

In other words, I say, when we face a conflict, especially when we cannot easily find an acceptable compromise, let's do exactly the same thing we do when we encounter a contradiction; let's insist that one of the underlying

assumptions is faulty. If, or should I say when, we pin down the underlying assumption that can be removed, we remove the cause of the conflict; we solve the conflict by eliminating it. (Used with permission by E. M. Goldratt, © E. M. Goldratt. All rights reserved.)

Up to this point, we have been discussing cause and effect in terms of “sufficiency.” (See [Fig. 25-9](#).) To say that “Y” is an effect of “X” is to say the following:

- If “X,” then we must have “Y.”
- “Y” exists because “X” exists.
- If “X” exists, then we know that “Y” must exist. If “Y” exists, “X” may not—something else might cause “Y” to exist.

When viewing cause and effect in terms of “necessity,” we are looking at conditions that must be in place in order for something (e.g., an objective) to be able to exist. To say that “B” is a necessary condition for “A” is to say the following (see [Fig. 25-10](#)):

- In order to have “A,” we must have “B.”
- We cannot have “A” unless “B” is in place.
- If we do not have “B,” then “A” is impossible.
- If “A” exists, we know that “B” must exist. However, if “B” exists, “A” may not—additional conditions may be necessary to cause it.

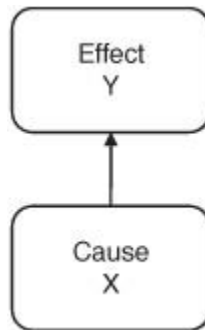


FIGURE 25-9 Sufficiency illustration.

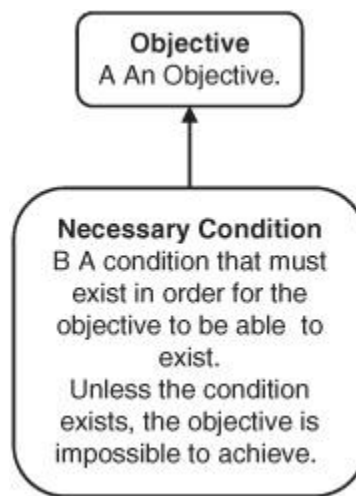


FIGURE 25-10 Necessity illustration.

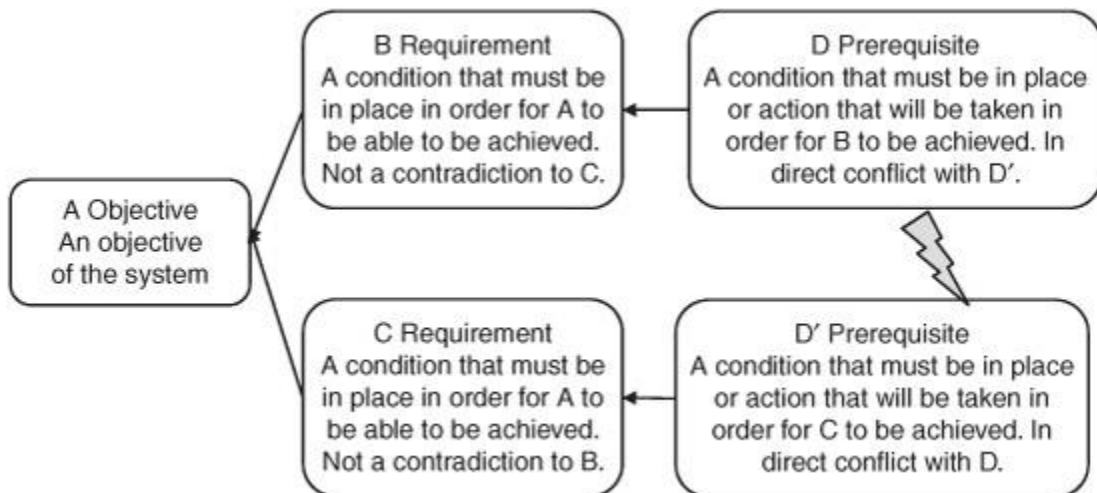


FIGURE 25-11 Cloud illustration.

The EC consists of five entities, and the arrows connecting them indicate the logic of necessity (see [Fig. 25-11](#)). The conflict itself—the conditions that are perceived as needed but that are in direct contradiction with each other—are described in the D and D' entities of the Cloud. “D” is a necessary condition for “B” and “D'” is a necessary condition for “C.” Both “B” and “C” are necessary conditions for “A.”

Once a Cloud is written, it provides several places for us to search for and locate the invalid assumption that is forcing the conflict—the perceived need for a contradiction (D and D'). If we could figure out that B is not really a necessary condition for A, then D is no longer necessary, and the conflict would be eliminated. Or, if we could figure out that D is not really a necessary condition for B, then it is no longer necessary, and the conflict could be eliminated. Or, if we could figure out that C is not really a necessary condition for A, then D' is not needed and the conflict would be eliminated. Or, if we could figure out that D' is not really a necessary condition for C, then it is no longer necessary and the conflict could be eliminated. Or, if we could figure out that D and D' are not really contradictions to each other and could actually coexist, then the conflict could be eliminated!

Necessity is not an established fact, but an interpretation.

—Friedrich Nietzsche

The Cloud is used to articulate any problem as a conflict, find the erroneous necessary condition relationship, and establish an injection that creates the path for a solution in which the conflict is fully eliminated. Some of the generic situations in which a Cloud is used are:

- Being caught between a rock and hard place—a decision needs to be made, and the only options available mean meeting the needs of one side and sacrificing the needs of the other.
- Eliminating gaps between authority and responsibility (the main cause for “firefighting” in organizations).
- Any argument between individuals, teams, organizations, and communities.

When TOC is implemented in operations, improving flow (reducing lead time) becomes an explicit, primary objective of the operation. Once the flow is put under control of the solutions such as *Drum-Buffer-Rope* (DBR) and *Buffer Management* (BM), the *Process of Ongoing Improvement* (POOGI) is put in place, in order to constantly improve the flow. The POOGI process for a *make-to-order* (MTO) manufacturer consists of documenting the answer to the question, “What is the order waiting for?” every time an order is delayed (not moving) for 10 percent of the production lead time. Periodically (e.g., weekly), a Pareto analysis is performed on the sources⁸ of all such delays that occurred for orders that the priority system (BM) indicated were at risk of becoming late. Teams are then put to the task of analyzing and eliminating the major sources of delay.⁹ The Cloud is a critical tool that teams use to analyze and solve the major source of delay. An example of this is provided as the steps to use an EC are described. You will also find detailed guidelines for using the Cloud on a day-to-day basis in [Chapter 24](#).

Make the Room Warmer	Make the Room Colder
Reduce the workforce	Don't reduce the workforce
Raise prices	Don't raise prices
Include government option in healthcare bill	Don't include government option in healthcare bill
Allow my teenager to stay out past midnight	Don't allow my teenager to stay out past midnight

TABLE 25-1 Examples of D and D' Conflicts

1. Write the D and D' entities of the Cloud. Write them in a way that it is obvious that they are mutually exclusive. Some examples are in [Table 25-1](#).

Our example company makes heavy steel products. In order to form and machine the steel to their customer specifications, the process includes heat treat—putting the product in large ovens to heat the steel and then placing the product in a tank of liquid (quench tank) to cool it rapidly and bring it to possess the metallurgical properties needed. The weekly POOGI Pareto analysis revealed that the most frequent answer to “What is the order waiting for?” was “Waiting for heat treat.” A POOGI team was assigned to analyze and eliminate heat treat as a major source of delay. As they reviewed the data, they found that the vast majority of the delays could be further classified as “green¹⁰ orders waiting for the assigned quench tank to become available.” They began to construct the Cloud ([Fig. 25-12](#)).

2. Write the corresponding B and C entities.

- B should answer the following questions:
- For what is D needed?
- What need will not be met if D doesn't materialize?

You should be able to fill in the blanks to the following statements:

- B won't happen without D.
- In order to have B, we must D.

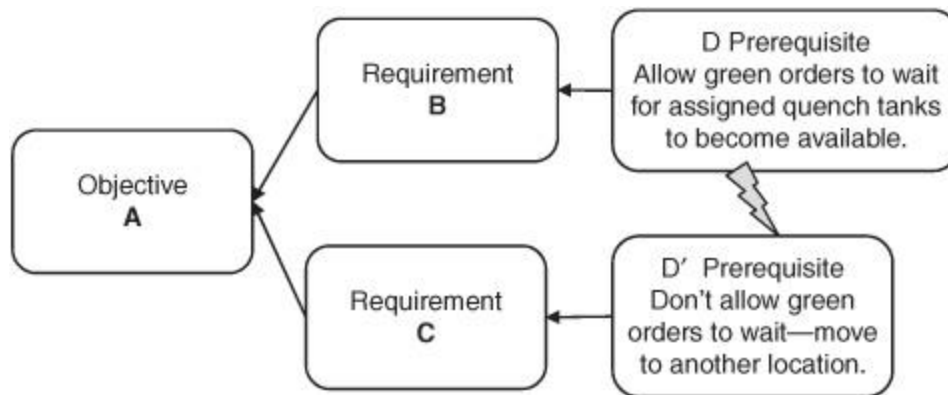


FIGURE 25-12 Cloud example 1.

- C should answer the following questions:
- For what is D' needed?
- What need will not be met if D' doesn't materialize?

You should be able to fill in the blanks to the following statements:

- C won't happen without D'.
- In order to have C, we must D'.
- The following check will also help:
- If D exists, then C cannot.
- If D' exists, then B cannot.

The POOGI team's analysis led them to understand the internal policy that forced orders to wait for quench tanks. It was not the lack of usable quench tanks in the company; rather it was the unavailability of the *specific* quench tank

defined in the order’s routing. The company had previously set a policy that allowed production managers to move orders to capable work centers other than those specifically identified in the routing when the priority system indicated that the order was becoming at risk of being late (yellow or red) or already late (black). In order to avoid “unnecessary expenditures” of time (making changes to paperwork) and money (transportation costs to move the product from one plant to another), the company did not allow such “exceptions” for “green” orders. Our steel products company’s Cloud now looked like the illustration in [Figs. 25-13](#) and [25-14](#).

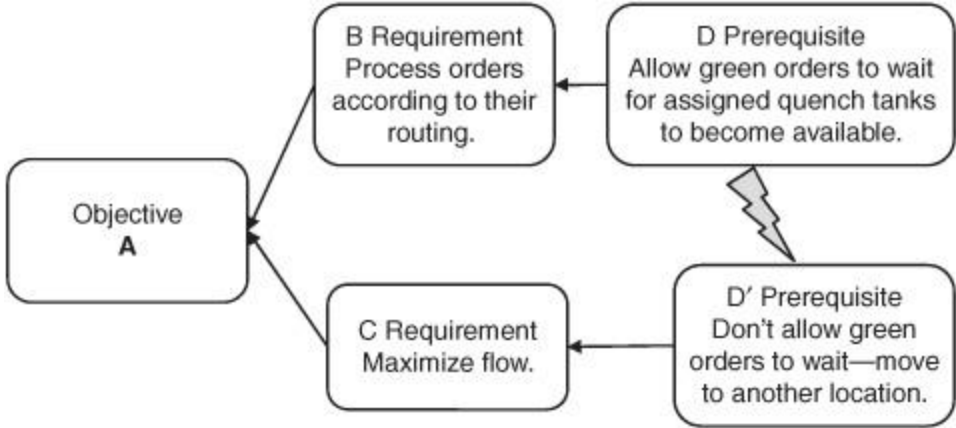


FIGURE 25-13 Cloud example 2.

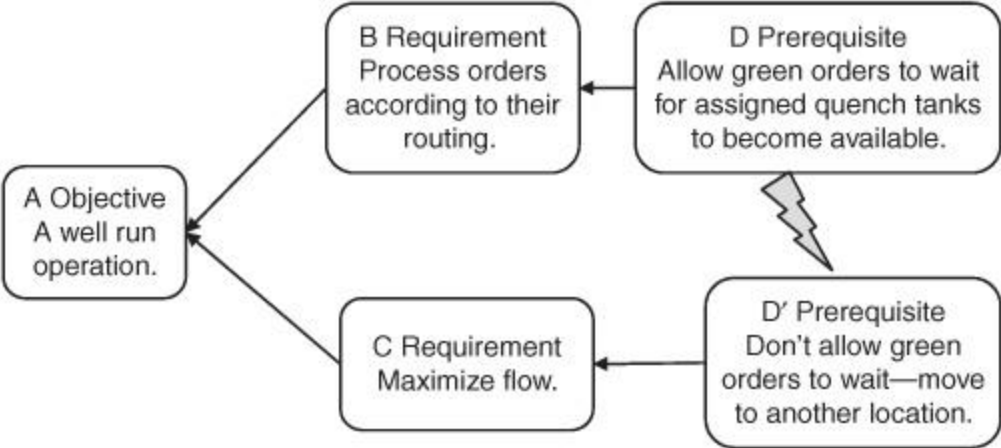


FIGURE 25-14 Cloud example 3.

Identify A, the mutual objective of B and C. Similar questions will enable you to verbalize the objective.

You should be able to fill in the blanks to the following statements:

- [A] won't happen without [B] and [C].
- In order to have [A], we must [B] and [C].

Our steel products company POOGI team completed their Cloud.

3. Surface the assumptions of each of the necessary condition relationships and identify those that are invalid in the situation of conflict being analyzed.

The Cloud (as well as the PRT) utilizes the logic of *necessary condition*. [Figure 25-15](#) illustrates the relationship between this logic and the logic of cause and effect that we have been using thus far.

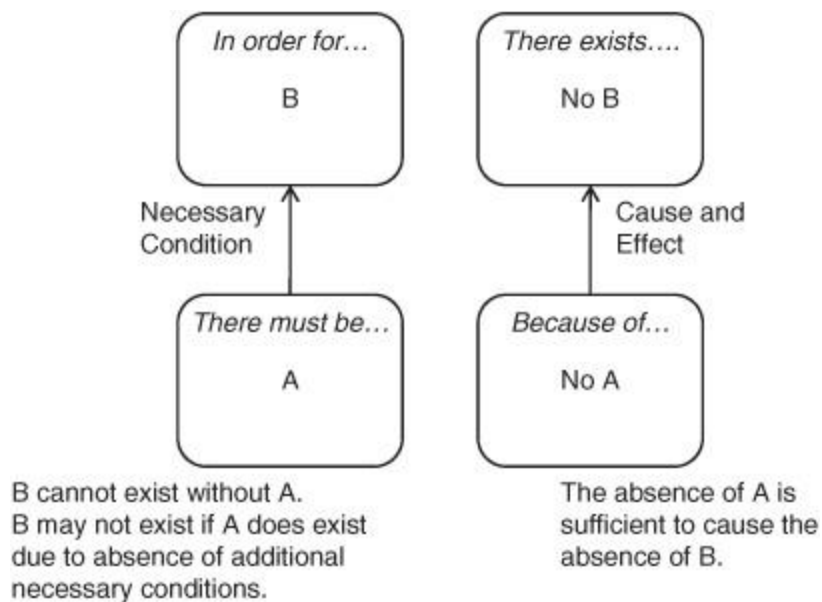


FIGURE 25-15 The relationship between necessary condition and cause-and-effect.

By understanding this relationship, you can surface—and check the validity of—the assumptions that are being made by using some simple questions and fill-in-the-blank statements:

- In order for A, we must¹¹ B, because _____.

Why can't A happen without B?

- In order for A, we must C, because _____.

Why can't A happen without C?

- In order for B, we must D, because _____.

Why can't B happen without D?

- In order for C, we must D', because _____.

Why can't C happen without D'?

- D and D' cannot coexist because _____.

Why can't B happen if D' exists?

Why can't C happen if D exists?

Note that you are looking for the “beliefs” that exist in the given situation. In [Table 25-2](#) some of the assumptions surfaced by the steel products company POOGI team are given.

4. Using the erroneous assumption as your guide, define an injection that would enable the conflict to be eliminated. A good injection will enable you to

“evaporate” at least one of the arrows in the Cloud. You should be able to fill in the blanks to at least one of the following sentences:

- If [injection], then [A] can be achieved without [B] because _____.
- If [injection], then [A] can be achieved without [C] because _____.
- If [injection], then [B] can be achieved without [D] because _____.
- If [injection], then [C] can be achieved without [D'] because _____.
- If [injection], then [D] and [D'] coexist because _____.

The analysis of the steel products POOGI team uncovered the following facts, which were in direct contradiction with existing policies:

- Allowing green orders to sit was not helping the company maximize flow, and in many cases led to expensive expediting later in the process.
- Moving an order to an equivalent resource that has open capacity, even if that resource is located at another nearby plant, is the most cost-effective approach to managing production.
- The routings had not kept up with the growth of the company—as equivalent resources had been added, the routings continued to identify a specific resource at a specific plant.
- As the company’s TOC implementation had progressed, the plant managers and supervisors of the various plants had established robust interplant communications, and it could be quite easy to identify where to move orders in order to ensure that orders

“sit” only when there is no capable resource available to process them.

The injections, then, became obvious and were communicated and implemented within days:

- If the resource on the routing is busy and another equivalent resource is available, move orders to any equivalent resource that is available, irrespective of color.

Necessary Condition Relationship	Assumptions
A←B (Why must we process orders according to their routings in order to have a well-run operation?)	... Because the routings identify the most appropriate way to process the order, given cost and quality considerations.
A←C (In order to have a well-run operation, we must maximize flow ...)	... Because maximizing flow enables us to reduce lead times and WIP and be more competitive in the marketplace.
B←D (Why do we need to allow green orders to wait for assigned quench tanks to become available in order to process orders according to their routings?)	... Because there are no other resources capable of doing the job at like cost and quality. THIS IS THE ASSUMPTION THAT THE TEAM REALIZED WAS INVALID.
C←D' (In order to maximize flow, we must not allow any orders to wait—we must move them to any available, capable resource ...)	... Because an order that is not moving is experiencing a delay in flow. ... a delay that occurs when an order is green may be the real cause for the order to become red later in its process.
D≠D' (Why can't D and D' coexist?)	... Because the routings specify a specific location, even though we have multiple tanks in other nearby plants that have the same capability. THIS ASSUMPTION WAS HELPFUL FOR THE TEAM TO FORMULATE THE SOLUTION.

TABLE 25-2 Steel Products Company Necessary Condition Assumptions

- Modify the routings so that equivalent resources are not an exception. (Upon subjecting the injections to NBR, the company decided to take the approach to modify routings as new orders are placed. As a *make-to-order* (MTO) company, this enabled the company to modify routings as they were needed, and avoided the expenditure of key personnel time on making unneeded modifications.)

If you would like to use the POOGI Cloud template in your organization, see Appendix D on the McGraw-Hill website: <http://www.mhprofessional.com/TOCHandbook>.

Conflict can be seen as a gift of energy, in which neither side loses and a new dance is created.

—Thomas Crum

The Integrated TOC Thinking Processes

The whole history of science has been the gradual realization that events do not happen in an arbitrary manner, but that they reflect a certain underlying order, which may or may not be divinely inspired.

—Stephen Hawking

We have explored the fundamental assumptions and basic building blocks of the TOC TP, in terms of the way that cause-and-effect logic, the protocol used for mapping the logic, the mindset required, and the scientific premise on which TOC and the TP are based. By putting to use the basics, you will be well prepared to use the full set of TP in order to *improve any system*.

To improve something means to make it better. And the only way for something to get better is if it changes. Think about the vast number of variables in any organization, relationship, or individual that *could* be better. If this is difficult to imagine, just think about the number of complaints you make or hear throughout any given day! If you agree that some improvements are better than others, and that the list of potential improvements outstrips the capacity available to make improvements, then you would conclude that in order to ensure a meaningful state of ongoing improvement, we must be able to systematically answer three fundamental questions (Goldratt 1990):

1. *What to Change?* Given everything that *could* be changed, what *should* be changed? No person or organization has infinite time on their hands, so if we are going to spend time making changes, it behooves us to distinguish between the important few and the trivial many. We should have a way to identify the variables that, if changed, could render the most significant improvement to the system.

Throughout this chapter, I use the words *system* and *situation*. I am not using them synonymously, though. A *system* is “a group of interacting, interrelated, or interdependent elements forming a complex whole.” A *situation* is “the combination of circumstances at a given moment; a state of affairs” (*The American Heritage*[®] *Dictionary*, 2004). We need both an understanding of the system itself and of the situation (state) in which the system finds itself, in order to find the answer to “What to change?”

2. *To What to Change?* Once we pinpoint what we want to improve, we should define the improvement itself—the future improved state we intend to create—and articulate the specific changes that need to be put in

place in order for the desired improvement to become the reality.

3. *How to Cause the Change?* By answering the first question, we have defined the critical few variables in the system that we intend to change in order to improve the situation. We have then designed the future improved scenario, highlighting the changes to make which will create the new reality. Now we need to draw the map and detail the action plan that, when followed, should bring us from the present to the improved future. The three questions of change are pictured in [Fig. 25-16](#).

The TOC TP are the tools used to answer the three questions of change. The *Current Reality Tree* (CRT) uses cause-and-effect logic to create a map of the existing situation and pinpoint a core problem—the common cause for many undesirable effects—and the answer to the question, “what to change.” With the EC, the problem is verbalized as a conflict, and a direction for a win-win solution is established by uncovering and replacing at least one erroneous assumption of the conflict. The *Future Reality Tree* (FRT) and NBR provide the process to create the logical model of the future system. They are used to answer the question, “to what to change,” highlighting the cause-and-effect relationships between the changes that will be made and the desired future state that those changes are intended to create. The PRT and *Transition Tree* (TRT) are the tools that TOC provides in order to logically derive and map what we need to do to close the gap between the current state and the desired future. With these tools, we clarify the obstacles that stand in our way, and what needs to happen in order to overcome them. The newest addition to the TOC TP—the *Strategy and Tactic Tree* (S&T)—provides for the full synchronization and communication of the

implementation of a change. In [Table 25-3](#), we see the purposes and relationships of the TP tools.

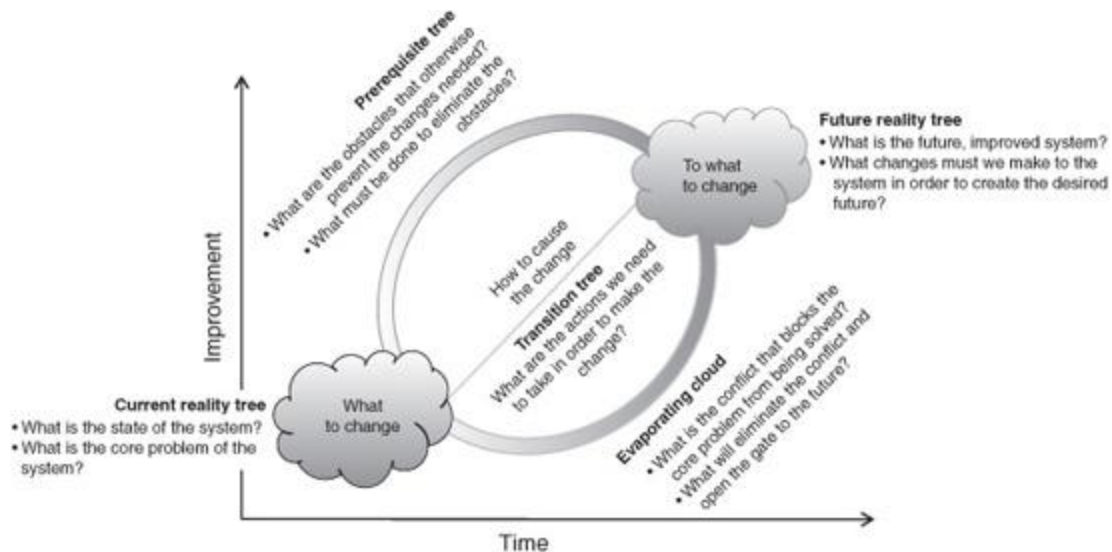


FIGURE 25-16 The three questions of change.

I am sure we are all guilty of having what we think is a great idea, and then falling in love with that idea to the extent that we spend our energy justifying, rather than validating, the value of the idea. A great way to *not* improve a situation is to fool yourself about what the situation really is and implement a solution for a non-problem. There is a term for this in TOC—*choopchick*. A Yiddish slang word originating in Serbia, a *choopchick* is generally translated as a *triviality*. In TOC, it is a dangerous form of triviality—it is a triviality that is believed to be important, and thus a distraction from what the focus of attention should be. By making the decision to take an internally honest, scientific, logical approach to answering the three questions of change, we can help avoid implementing non-solutions and chasing choopchicks.

Thinking Processes	What to Change?	To What to Change?	How to Cause the Change?
Current Reality Tree	Core Problem		
Evaporating Cloud	Core Conflict	Breakthrough Injection	
Future Reality Tree		Solution	
Negative Branch			
Prerequisite Tree			Intermediate Objectives
Transition Tree			Actions
Strategy and Tactic Tree	Communication and Synchronization		

TABLE 25-3 The Purposes and Relationships of the TP Tools

The effect of choopchiks within the management process can be devastating. Attracting attention to relatively unimportant issues diverts efforts from genuinely significant concerns.

—John Caspari, *Handbook of Management Accounting*

Reinforcing the Mentality of a Scientist—Jonah’s Approach

It is one thing to get on my soapbox and ask you to be internally honest, scientific, and logical. However, this chapter is about providing you with a practical means to actually do so. Here are four simple steps that can guide you to a good understanding of the present situation, the future you want to create, and the decisions and actions you would need to take to turn the future you want into reality.¹²

1. **Entity Existence.** Verify that each entity really does exist in the environment that is being analyzed. If an entity is something that cannot be directly confirmed, physically observed, or numerically verified—use the

scientific method. For instance, a person smiling is something that is physically observed. What a person is thinking, or what we assume is a person's attitude, is not physically observed and can only be directly confirmed by the person. Predict another effect that must exist as a result, and check for it. If the effect exists, you have increased the likelihood that the intangible effect exists. If the predicted effect does not exist, then you have eliminated the likelihood that the intangible effect exists.

Let us revisit the lamp example from earlier in the chapter. At one point, we predicted that the power was out in the rest of the neighborhood. The street was dark, which was an additional effect of a neighborhood power outage. If we had looked outside and saw all of the streetlights and the lights in our neighbors' homes brightly lit, then we would have known that there was not a power outage in the neighborhood. It would *not* have been an entity that existed in the situation we were analyzing.

2. Entity Clarity. Ensure each entity is stated clearly and concisely, as a simple yet complete sentence. A good test is that when you read the entity statement aloud, it needs no further explanation. An indicator that the statement is not yet clear enough is if you read it aloud to someone and feel compelled to explain further what it means.

3. Causality Existence. Validate that each cause and effect relationship identified in the analysis really does exist in the situation being analyzed. Even when you verify that the described elements do in fact exist in the situation or system being analyzed, it could very well be that the hypothesized cause-effect

relationship between them does not. Here is an example.

I know a young woman who had a persistent headache. The headache was there when she woke up in the morning, throughout the day, and when she went to sleep at night. It simply did not go away. After a couple of weeks, she went to a local urgent care center.¹³ After asking a few questions and short examination, the doctor formulated his hypothesis and prescribed a solution accordingly. His hypothesis of the woman's problem was that she had a simple tension headache. He prescribed a painkiller and told her to go home and relax. A simple analysis of the situation, in the doctor's view, would have looked like Fig. 27-17a. Unfortunately, even though every entity in the tree did exist, and even though for most young adults stress is the cause for a headache, it was not in this case.

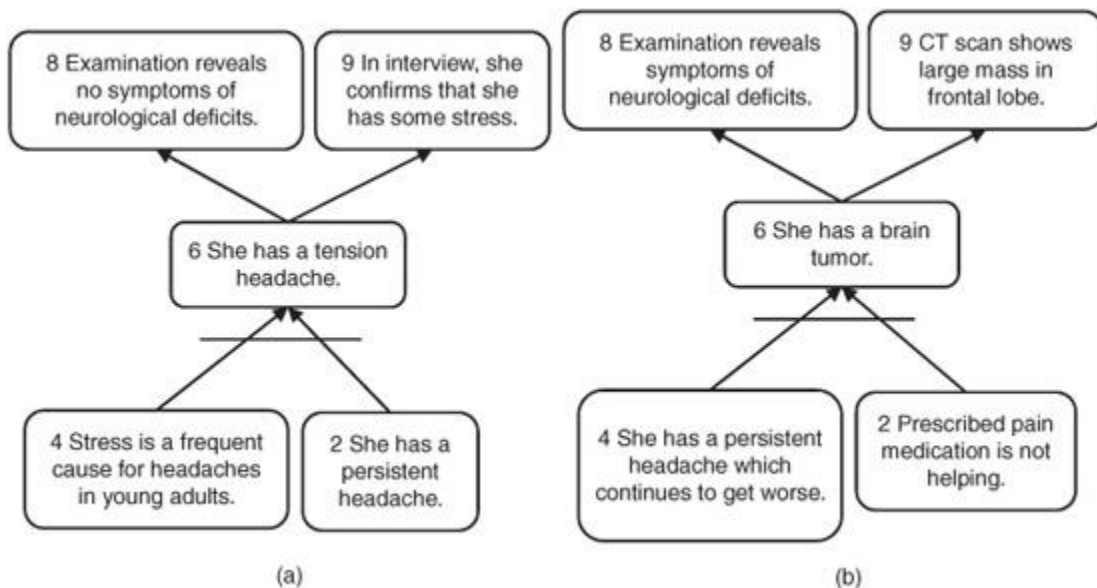


FIGURE 25-17 Validating “causality existence.”

A week or so and many pain pills later, the headache was not only still present, it had worsened, and she had become nauseated and disoriented. The young woman went to the emergency room at a local hospital. After a short interview and examination, the ER doctor formulated his hypothesis, which was that there was something physically going on in her head, possibly a tumor. He ordered a CT scan, which verified the existence of a quite large tumor in the left frontal lobe of her brain. (See [Fig. 25-17b.](#))

I *am not* illustrating this case in order to pass judgment on either of the two doctors involved. I *am* illustrating this case in order to show that even though the same conditions might exist in two different realities, they have a cause-and-effect relationship in one of those realities and not another. The young woman did have some stress in her life, and she did have a headache. Tension is the cause for headaches often, but not always.¹⁴

Check the causality! It doesn't take long to ask any or all of these questions: Why? ♦ How do I/we/you know? ♦ Is this always the case? ♦ Under what circumstances is this the case? ♦ Under what circumstances is this not the case? ♦ Oh, really? ♦ Why?

4. Causality Clarity. Ensure each cause-effect relationship is modeled clearly and concisely. A good test is to read aloud the relationship as an “if-then” statement or as a “because” statement. An indicator that the cause-and-effect relationship is not yet clear enough is if you read it aloud to someone and feel compelled to explain further what it means. For instance, look at [Fig. 25-18](#). The cause-and-effect relationships would be read as:

- If [B] and [C], then [A]; or [A] exists because [B] and [C].
- Additionally, if [D], then [A]; or [A] also exists because [D].

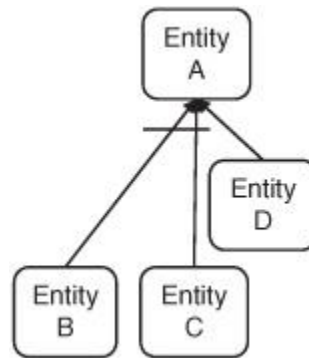


FIGURE 25-18 Cause clarity.

As we explore the full, integrated TOC TP, I will use examples from the case study of a bank, which was described in detail by Cox, Blackstone, and Schleier (2003) in their book, *Managing Operations: A Focus on Excellence*. (Used by permission, © Cox, Blackstone, and Scleier)

What to Change?

Mad let us grant him them, and now remains that we find out the cause of this effect—Or rather say, the cause of this defect, for this effect defective comes by cause. Thus it remains, and the remainder thus.

—William Shakespeare

In order to answer the question, What to Change?, we will use two of the TP tools: the CRT and the EC. Over the years, two approaches have emerged as “standard.” The “Snowflake Method” is considered to be the more traditional approach, mainly because it is an older method than the

“Three-Cloud Method,” and the “Three-Cloud Method” is generally easier for people to learn. The main difference between the two approaches is the sequence in which the two tools are used, and in which the core problem is identified. The “Three-Cloud Approach” tends to be easier to learn. Both methods have proven to be quite effective in gaining an understanding of the situation and the core conflict (core problem) that has prevented the otherwise natural harmony to be in place.

Current Reality Tree (CRT)

We find in the course of nature that though the effects be many, the principles from which they arise are commonly few and simple, and that it is the sign of an unskilled naturalist to have recourse to a different quality in order to explain every different operation.

—David Hume

A CRT is a cause-effect model of an existing situation. The main use of a CRT is to answer the question, *What to Change?*, so the cause-effect relationships that are focused on in the CRT are the UDEs—the aspects of the situation that we want to improve.

One important aspect of the inherent simplicity concept is *convergence*. Goldratt explains that “science is simply the method we use to try and postulate a minimum set of assumptions that can explain, through a straightforward logical derivation, the existence of many phenomena of nature” (Goldratt and Cox 1986, Introduction). When we look at a well-constructed CRT, we are able to see clearly the very few causes for a much larger set of effects.

The grand aim of all science is to cover the greatest number of empirical facts by logical deduction from the smallest number of hypotheses or axioms.

—Albert Einstein

Evaporating Cloud (EC)

The peak efficiency of knowledge and strategy is to make conflict unnecessary.

—Sun Tzu

By definition, a problem is something that we want to solve. In other words, if I have a problem, then I want to replace it with its opposite non-problem. Whether a given problem is a core problem (the cause for many UDEs), or an UDE (an element of the system that is undesirable), it is an obstacle to harmony that should be eliminated. This means that *any* problem can be verbalized as a conflict, which leads us to the use of the EC. In the “Snowflake Method,” the EC is used to summarize a core problem reflected in a CRT that has been constructed by logically connecting the UDEs. In the “Three-Cloud Method,” the Cloud is used to derive the core problem and then logically connect it with the UDEs.

The “Snowflake Method”

1. Pick a subject matter. What is the system or situation that you want to understand better in order to improve it? Perhaps you want to understand your markets better to develop a product or offer that would address a significant need; or you want to understand your organization better to determine why

it is not growing faster, serving its customers better, or retaining its employees longer; or you want to understand your supply chain to find the keys to improving the relationships with both your suppliers and your customers; or you want to understand your family or other relationships better to figure out what to change to make them more meaningful. Hospitals have used the TP to understand what needed to change to improve their emergency rooms and surgical centers; even a religious denomination¹⁵ used the CRT to understand what was preventing it from better accomplishing its mission. The list of potential subjects is limitless. There are two criteria that you should use to determine a subject on which you will construct a CRT:

- a. You really care about it to the degree that you intend to roll up your own shirt sleeves when it comes to implementing the solution.
 - b. You have enough experience to have some intuition about it.
2. Identify several aspects of the situation that are undesirable, and write them as entities. These entities are called UDES. *An UDE is defined as an entity that describes an element of the situation that we want to improve; in other words, it describes an aspect of the system that is undesirable, and which we would like to change.* Try not to identify fewer than six or more than twelve in this early step of the process. This simply defines the starting point for the analysis.
 3. Your intuition will point you to some of the UDEs that are closely connected to each other through cause and effect. Starting with these, construct the cause-effect map that shows how they are ultimately connected. Remember to verify that the entities really

do describe elements of the situation as it exists, validate the causality, and ensure that what is written is clear and understandable. Once you are satisfied that you have a cluster that is solid from a logical cause-effect perspective, go back to your list of UDEs and, one by one, let your intuition guide you to the area on the tree to which they are connected, and then use cause-effect logic to connect them. Do not stop until all of the UDEs are contained in the diagram.

4. If your intuition tells you that the tree you have is not telling the whole story, add the causes and effects so that it does. You may also discover that many of the entities you initially defined as UDEs really are not, but that others in the tree really are. Go ahead and identify the “*real* UDE’s.”¹⁶ Remember to keep the view of the scientist.
5. Check the entities that are causes only. Can you identify one that is responsible for the majority (say 70 percent or more) of the UDEs in the tree? If so, you have uncovered a core problem. If not, select the few that together are responsible for most of the UDEs and see if you can identify the common cause for them. If not, don’t worry—your work on the CRT has provided you with enough understanding of the situation that you would be able to use an EC to clarify the core problem and establish a direction for the solution.
6. Construct the EC in order to crystallize the core conflict of the system. There are two approaches to constructing the Cloud from a CRT. One approach is to summarize the CRT. Another is to use the core problem that has been identified in the CRT as the D

entity, its opposite as the D' entity, the goal of the system as the A entity, and fill in B and C based on the understanding of the system that has been established by constructing the CRT.

The Bank Case: What to Change, Snowflake Approach

A brief background (step 1) to the bank case, as provided by Cox et al. (2003):

The bank has a problem with employee turnover and pay levels. Other businesses pay more than the bank can pay for entry-level positions and hire the bank's employees. Employees are constantly turning over so the bank is unable to develop loyalty with its customers

. . . .

In order to get a holistic view of the bank and not just of an individual within the bank, the Branch Manager, the Head Cashier, and the Director of Human Resources defined the UDEs (step 2). They checked for the existence and clarity of the entities, and after some wordsmithing, the list of UDEs they used to begin their CRT was:

1. Many bank tellers quit and take better job positions.
2. Some single-parent bank tellers quit to make more money on public assistance and be with their children.
3. Many bank teller job vacancies occur each year.
4. The bank's budget for hiring, training, and raises is quite small.
5. Some bank tellers (students or their spouses) quit at college graduation.

6. Bank teller jobs are low paying entry-level positions.
7. The bank loses a lot of revenue from past, existing, and potential customers.
8. Some tellers make errors in customer accounts.
9. Some tellers do not know how to handle multiple complex transactions.
10. Some tellers are extremely slow.
11. Many customers go elsewhere to bank.
12. Many customers complain about poor service to other customers (existing and potential).
13. New employees do not know the names, likes, and dislikes of loyal customers.

The team immediately identified three causes for UDE #3 and mapped them accordingly (step 3), as illustrated in [Fig. 25-19](#).

They then added UDE #6 to the cluster (step 3-2 [Fig. 25-20](#)).

They continued to follow the steps (step 4), and [Fig. 25-21](#) is the CRT on which they agreed reflected the reality of the situation.

of the vast majority of the “real UDEs.”

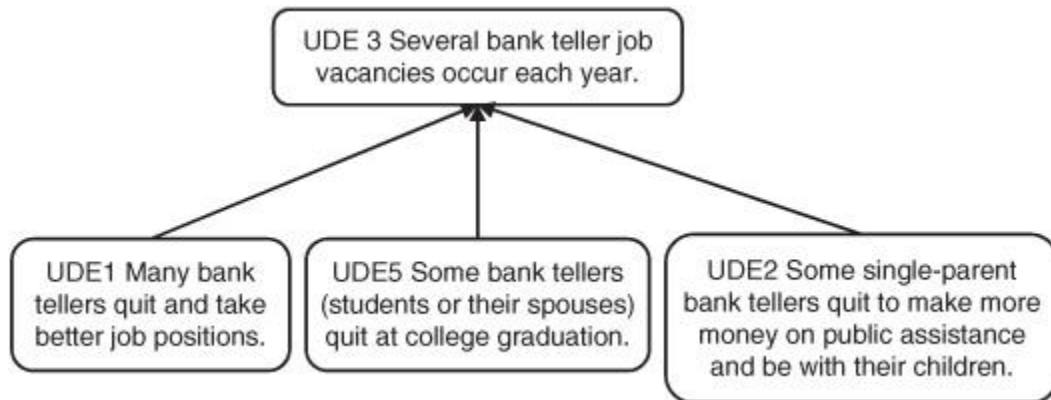


FIGURE 25-19 Bank CRT step 3.

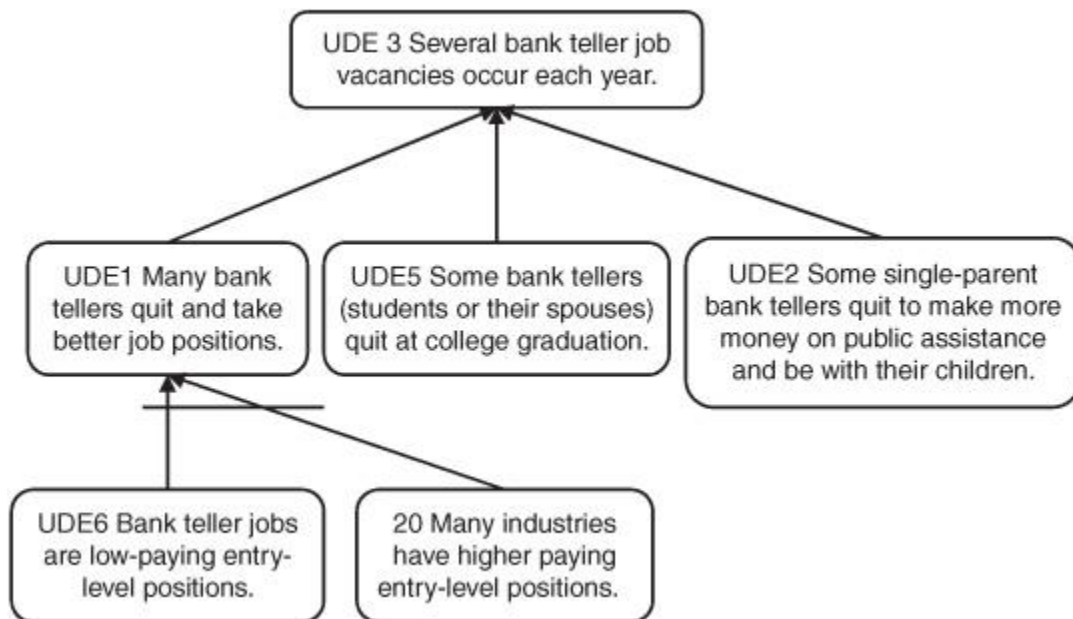


FIGURE 25-20 Bank CRT step 3-2.

As you examine the bank’s CRT, you may find yourself questioning some of the entities and the cause-effect relationships as they are represented in the model. If so, and if you had been sitting in the room with the bankers at the time, your reservations might have helped them end up with a more “perfect” CRT. Nevertheless, I do believe this is a “perfect example” to share with you. It is from real life, not an ivory tower. Real managers expended real human

energy to understand their environment better for the purpose of making decisions and taking actions that would cause real improvement for their bank and their customers. “Perfect” logic may be a good aspiration to help you keep the mindset of the scientist. However, it is quite inappropriate to spend an exorbitant amount of time to map out “the perfect CRT.” Do not allow “analysis paralysis” to set in! As you will see, the full set of TP provides excellent safety nets. Even if the CRT is not “perfect,” the subsequent steps will help you pick up anything important that you may have missed.

The Branch Manager summarized the CRT in the Cloud shown in [Fig. 25-22](#).

The bank team identified entity #140 (step 5), “The bank is unable to maintain an adequate pay structure to provide stable employment.” If the bank would have instead constructed the Cloud using the core problem entity as the D entity of the cloud, the Cloud may have looked like the one shown in [Fig. 25-23](#).

Note that in either case, the conflict is well represented in the CRT (step 6).

subsequent steps are used to identify the cause-effect connections between the core conflict and the UDEs. We will pick up from Step 3.

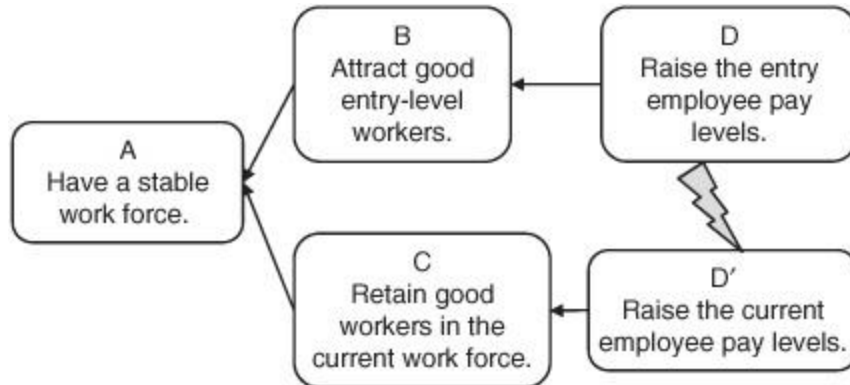


FIGURE 25-22 Bank Cloud.

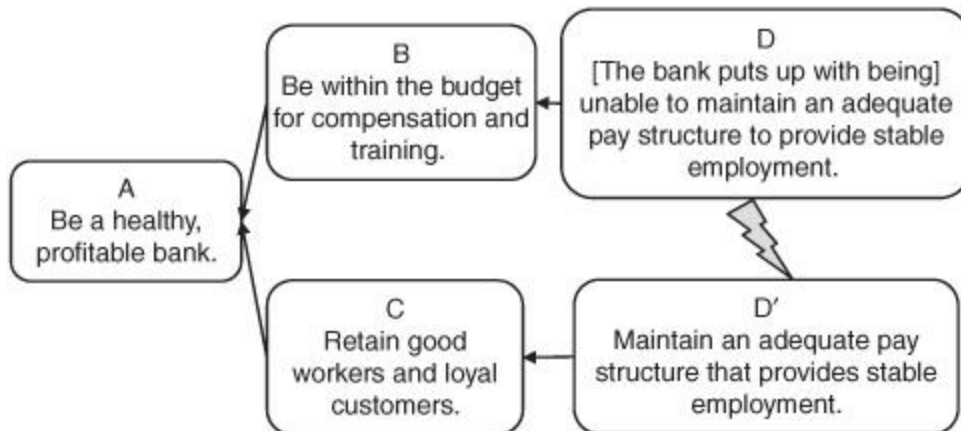


FIGURE 25-23 Bank Cloud 2 UDE the bank puts up with being unable to maintain an adequate pay structure to provide stable employment.

3. Select three UDEs, making sure to select them from diverse aspects of the system. A good guideline to follow is to select UDEs that do not seem to be connected to each other via cause and effect. Create

a Cloud for each of the selected UDEs according to the template shown in Fig. 15-24.

Three of the bank's UDEs, verbalized as ECs, are shown in [Figs. 25-25](#) through 25-27.

4. From the three Clouds, create the Generic Cloud of the system, which is the core conflict. When you examine the three Clouds together, you will be able to uncover a theme for the As, the Bs, the Cs, the Ds, and the D's. I find [Table 25-4](#) useful, and have used it to illustrate how the bank's three specific UDE Clouds are converted into a Generic Cloud.

Now you can create the Generic Cloud as seen in [Fig. 25-28](#).

The bank's Generic Cloud, according to the Three-Cloud Method, is shown in [Fig. 25-29](#).

Notice the similarity between the Cloud in [Fig. 25-29](#) and the Cloud that was generated with the core problem (entity #140) identified in the Snowflake Method ([Fig. 25-22](#)).

5. The CRT is completed by establishing the cause-and-effect linkages between the core problem and the UDEs.

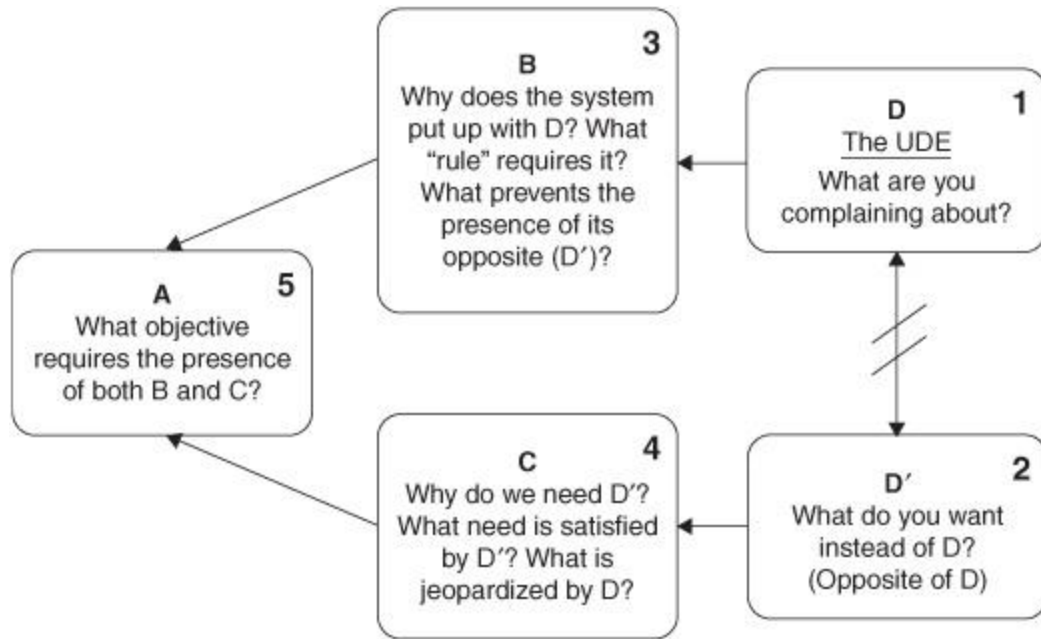


FIGURE 25-24 Template for UDE Clouds.

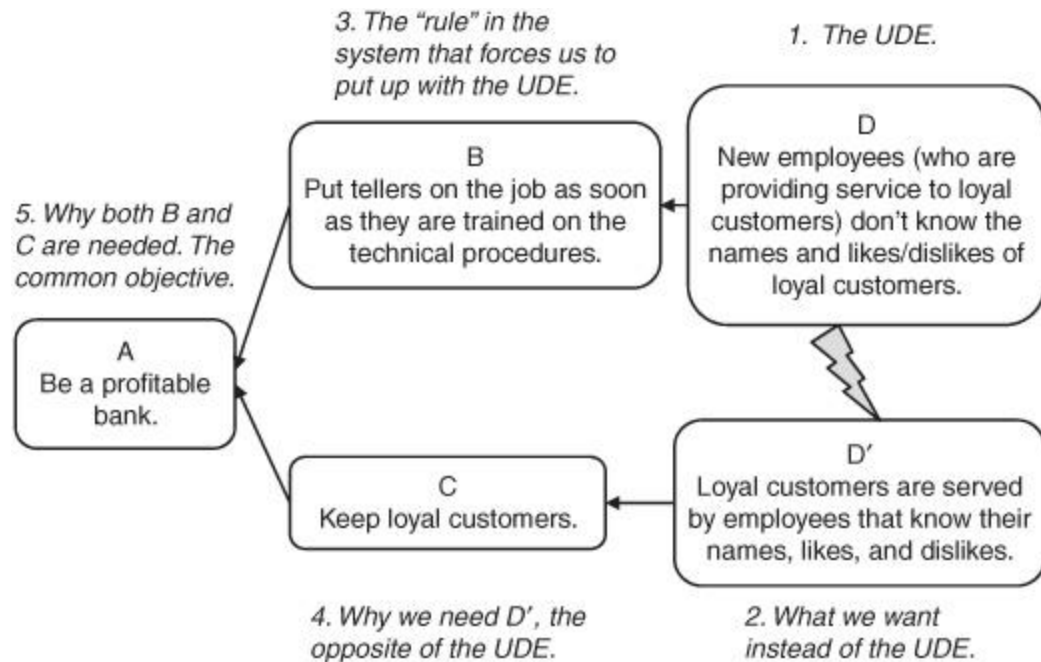


FIGURE 25-25 EC for bank UDE 13.

To What to Change

My interest is in the future because I am going to spend the rest of my life there.

—Charles F. Kettering

We will utilize a few of the TP tools to answer the question, To What to Change? The Cloud that has already been constructed is used to surface assumptions, identify those that are invalid, and define the initial injection for the solution. We will then complete the solution with the FRT and NBR.

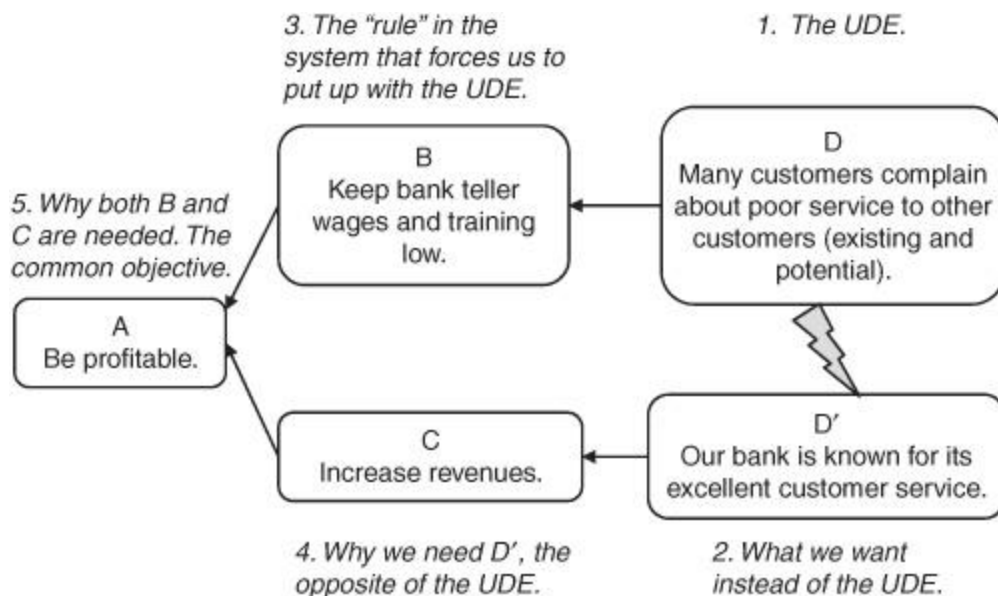


FIGURE 25-26 EC for bank UDE 12.

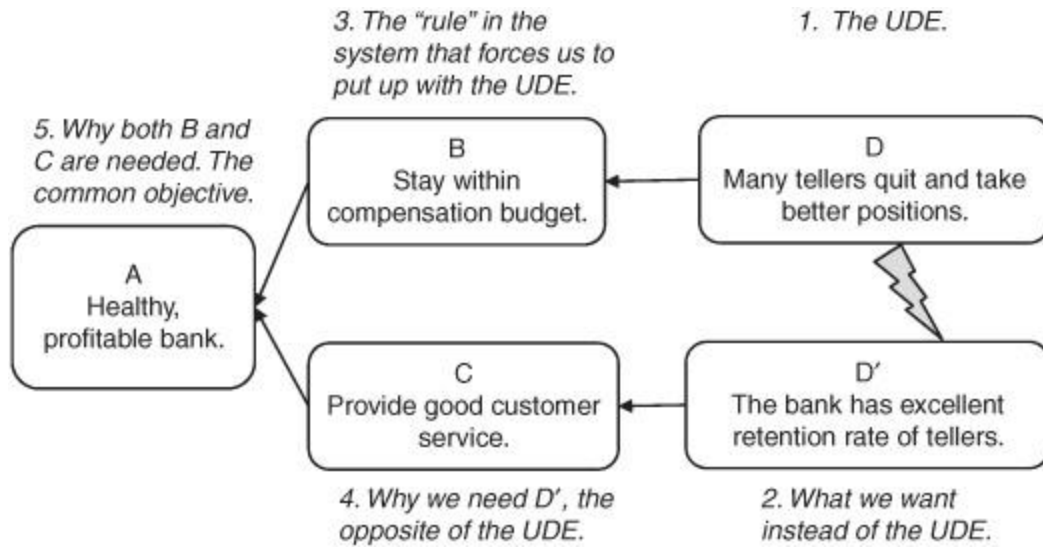


FIGURE 25-27 EC for bank UDE 1.

A	A from first Cloud: Be a profitable bank.	Generic A: Be a healthy, profitable bank.
	A from second Cloud: Be profitable.	
	A from third Cloud: Healthy, profitable bank.	
B	B from first Cloud: Put tellers on the job as soon as they are trained on the technical procedures.	Generic B: Minimize costs (low pay and minimal training for tellers).
	B from second Cloud: Keep bank teller wages and training low.	
	B from third Cloud: Stay within compensation budget.	

C	C from first Cloud: Keep loyal customers.	Generic C: Maximize revenues (customer retention and loyalty).
	C from second Cloud: Increase revenues.	
	C from third Cloud: Provide good customer service.	
D	D from first Cloud: New employees (who are providing service to loyal customers) do not know the names, likes, and dislikes of loyal customers.	Generic D: Sacrifice the quality of service.
	D from second Cloud: Many customers complain about poor service to other customers (existing and potential).	
	D from third Cloud: Many tellers quit and take better positions.	
D'	D' from first Cloud: Loyal customers are served by employees who know their names, likes, and dislikes.	Generic D': Do not sacrifice the quality of service.
	D' from second Cloud: Our bank is known for its excellent customer service.	
	D' from third Cloud: The bank has excellent retention rate of tellers.	

TABLE 25-4 Converting the Bank’s Individual UDE Clouds to a Generic Cloud

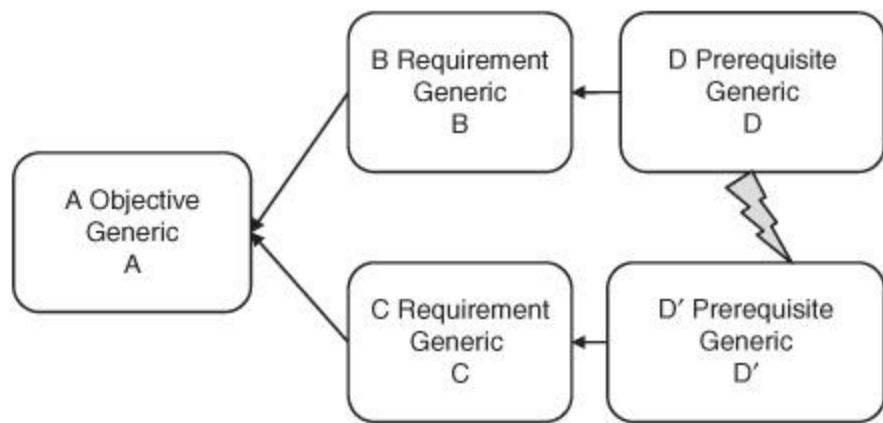


FIGURE 25-28 Generic Cloud template.

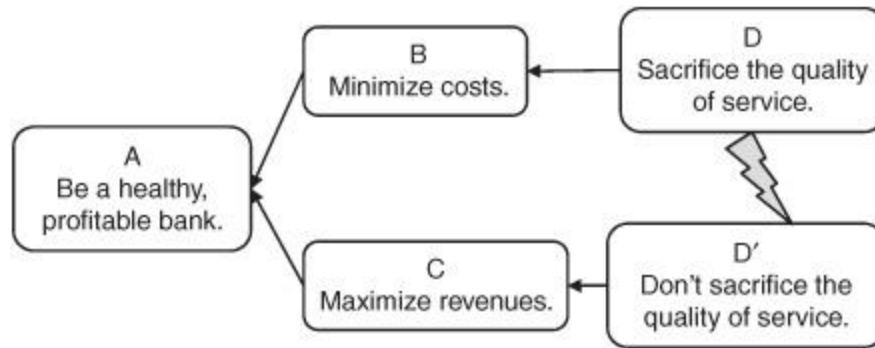


FIGURE 25-29 Bank Generic Cloud based on three clouds.

Evaporating Cloud

There are three ways of dealing with difference: domination, compromise, and integration. By domination only one side gets what it wants; by compromise neither side gets what it wants; by integration we find a way by which both sides may get what they wish.

—Mary Parker Follett

Earlier in the chapter, as well as in [Chapter 24](#), we learned how to surface assumptions and identify injections using the Cloud. Therefore, let us go directly to the bank case. The bankers used the Snowflake Approach to build their CRT and the Cloud they used was the summary Cloud ([Fig. 25-22](#)). The team examined the various necessary condition relationships, and when they reached the assumption that held D and D' as being in contradiction with each other, they realized that they had found the key to the solution. The reason that the bank was unable to raise the pay levels of entry-level employees and raise the pay levels of existing employees was that *the bank's budget for hiring, training, and raises couldn't be increased*. Nobody at the bank had

the authority to increase the total budget for hiring, training, and raises. However, the branch manager did have authority over the total budget. What would happen if they were able to shift money from hiring and training to salaries? If such a shift could enable the bank to pay new employees more, and also enable the bank to better reward existing employees, then the turnover would be reduced, and the volume (and thus the cost) of hiring and training would be reduced!

The injection the bank used to begin to develop its solution was, “The bank uses monies for hiring and initial training to raise the pay for entry position pay levels.” We now use this initial injection as the starting point for the full solution that will be detailed in the FRT.

Future Reality Tree and Negative Branch Reservation

A human being fashions his consequences as surely as he fashions his goods or his dwelling. Nothing that he says, thinks or does is without consequences.

—Norman Cousins

The FRT and the NBR are both processes that model the predicted effects of injections. The FRT is used to model the intended effects—the desired improvements—that comprise the full solution. FRTs typically contain several injections and many entities. They show the cause-and-effect model of how the injections enable the achievement of the objective of the Cloud and the opposite of (elimination of) the UDEs that were described in the CRT. The NBR is used to show how an injection would lead to undesired consequences, and then modify the idea (by modifying an injection or adding additional injections) to the degree that predicted undesirable consequences would be prevented. The

guideline is to build the FRT first, and then use the NBR process to modify and solidify the solution to ensure that it is win-win-win.

The steps to construct an FRT and NBR are shown in [Table 25-5](#).

The FRT (inclusive of the resolved NBRs) of the bank is shown in [Fig. 25-30](#). As with the CRT, when you examine the bank's FRT, I have no doubt you will identify entities that could use more explanation and causal connections that are flawed. Moreover, I have no doubt that if you had been with the team that constructed the tree, your reservations would have helped them to create a "more perfect FRT." Nevertheless, you are looking at real work done by real people. The results spoke for themselves. This analysis was completed 15 years ago. The bank implemented the injections. Employee turnover dropped like a stone, customer service improved, and the bank grew. A decade later, tellers and managers alike greeted customers by name, and the bank enjoyed the loyalty of its customers and employees.

Step	Future Reality Tree	Negative Branch Reservation
1	Write the initial injection. If the initial injection encompasses multiple elements, try to write each element as a separate injection.	Write the idea as an entity (injection). If there are multiple elements of the idea, try to write each element as a separate entity (injection).
2	Make a list of the intended benefits (effects) of the injection, and write each as an entity. The list should include the objective of the Cloud and the intended replacement for each UDE in the CRT.	Make a list of the pros (benefits) and cons (concerns) of the idea. Write the negative outcomes that you are predicting as entities—these are the predicted UDEs. Again, try to write each element as a separate entity.
3	Using the mapping protocol discussed earlier in the chapter, connect the injection to the intended benefits. Add injections as needed in order to complete the tree.	Using the mapping protocol discussed earlier in this chapter, connect the injection entity (or entities) to the predicted UDEs.
4	Use the guidelines for checking the validity of the cause-and-effect relationship discussed earlier in this chapter to scrutinize the FRT, and make adjustments accordingly. Where possible, reinforce the solution with positive causal loops.	Use the guidelines for checking the validity of the cause-and-effect relationship discussed earlier in this chapter to scrutinize the NBR, and make adjustments so that it reflects your full hypothesis.
5	Scrutinize the tree and check for NBRs.	Trim the negative branch. Identify the place in the tree where the transition from "neutral" to "negative" occurs. Identify a new injection or modification to an existing injection that, if implemented, would either prevent the UDE from occurring or replace the predicted UDE with an entity that would be an additional benefit of the solution. Check to make sure that this new, added injection does not lead to more negative ramifications. If it does, either replace it with a different injection or add an additional injection to trim the new negative branch.

TABLE 25-5 Constructing an FRT and NBR

Unfortunately, a few years ago the bank ended up being acquired by a larger bank, and then again by an even larger bank, and the policies and procedures of the conglomerates were installed. Neither tellers nor managers know the customers, and rarely does one see a smile in the bank. Customer and employee turnover is back to the levels it experienced at the time of the original analysis.

The fish stinks from the head.

—Yiddish proverb

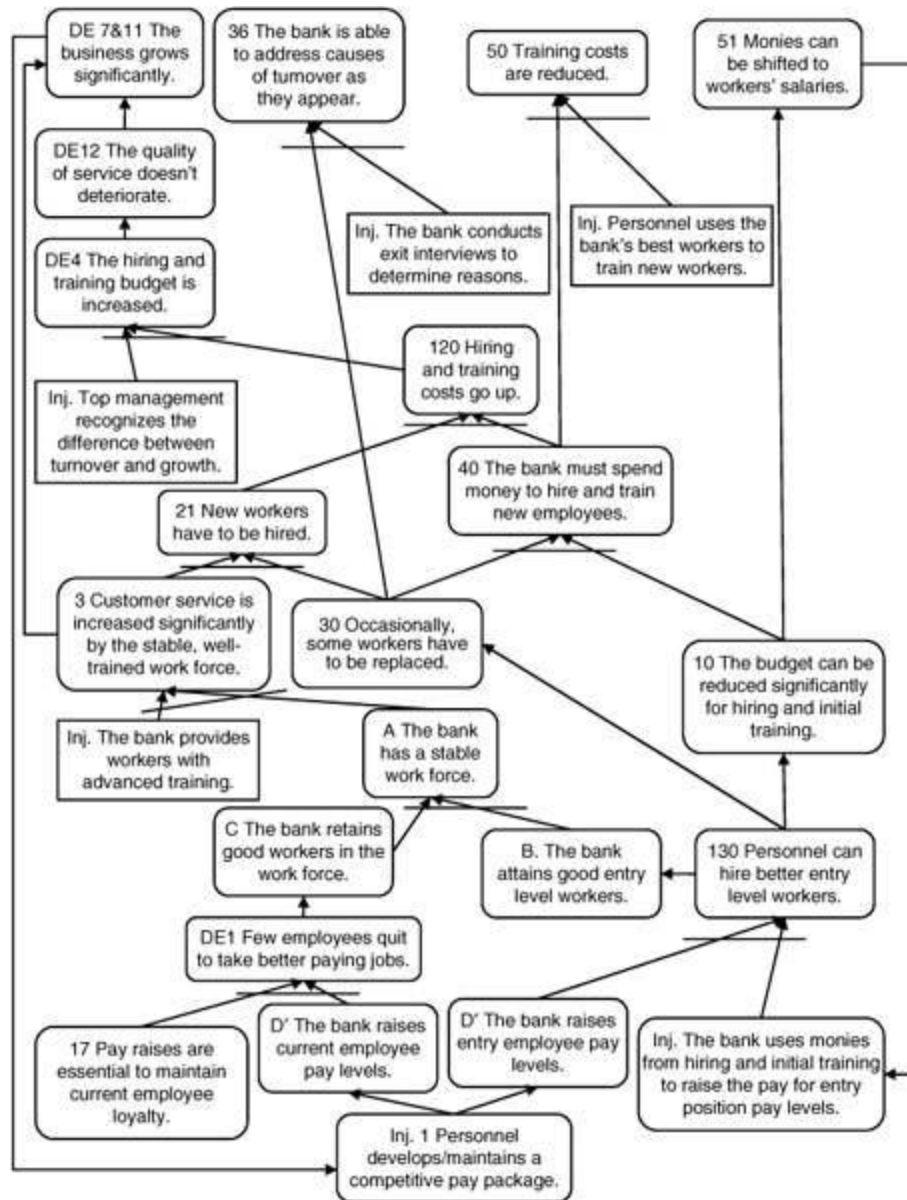


FIGURE 25-30 Bank FRT.

How to Cause the Change

A thought which does not result in an action is nothing much, and an action which does not proceed from a thought is nothing at all.

—Georges Bernanos

Three TOC TP are used to answer the third question of change, How to Cause the Change? With the *PRT*, we identify the obstacles that make implementation of the injections difficult and create a logical map of Intermediate Objectives (IOs) that will overcome the obstacles. *TRT* are used when it is necessary to define the specific, detailed actions that will be taken in order to achieve a given objective. Finally, the *S&T* tree is used to integrate the output of all of the TP into a synchronized whole that fosters communication and synchronization necessary for the successful implementation of major initiatives.

Prerequisite Tree

Obstacles don't have to stop you. If you run into a wall, don't turn around and give up. Figure out how to climb it, go through it, or work around it.

—Michael Jordan

The [PRT¹⁷](#) takes advantage of the same type of “necessity” logic approach as the EC. With the EC, we are modeling a set of necessary conditions that are thought to exist in the current reality of a conflict. With the PRT, we are building the necessary conditions to create a logical roadmap to move from the current situation to the desired future. We will use [Fig. 25-15](#) (which was used previously in the EC section of this chapter) to highlight the use of the logic. In both cases (Cloud and PRT), B cannot be achieved unless A is in place because of an aspect of the current reality that exists. When we are using the EC, we call this aspect of

current reality an *assumption*. When we are using the PRT, we call this aspect of current reality an *obstacle*. When we use the EC, we begin with the entities “in the boxes” (A, B, C, D, and D’), and then surface the assumptions. When we are using the PRT, we begin with the obstacles, and then define the entities “in the boxes” (intermediate objectives).

In [Chapter 24](#), you will find detailed instructions for creating a PRT. Here are the basic steps:

1. For each injection, list the major obstacles to achieving it. *An obstacle is an entity that exists in the current reality, which, due to the fact that it exists, prevents an injection from being reality.*
2. For each obstacle, define an IO—an entity that, once implemented, causes the obstacle to be overcome. An obstacle can be overcome by eliminating the entity or by finding a way around the entity (the entity would still exist; it would simply no longer be an obstacle to achieving the injection).
3. Using necessary condition logic, map the order in which the IOs must be implemented. [18](#)

The Bank’s Prerequisite Tree

The bank identified six injections in its FRT:

No.	Obstacles	Intermediate Objectives
21	The bank does not know what the wage rate structure is in their locality (banks and other industries).	Contact the local Chamber of Commerce to get most recent salary surveys.
22	The bank does not know what other banks are paying experienced help.	Bank checks with local banking association to determine pay levels of experienced help.
23	Management lacks a clear definition of what "competitive" means. What does it take to keep an excellent employee?	Management reviews above data and sets pay 10 percent above comparable work and pay.
24	Management does not have a clear policy for removing average to marginal employees.	Management establishes performance guidelines and measures providing feedback to employees. Marginal to average employees are terminated.

TABLE 25-6 Obstacles and Intermediate Objectives for the I/O Map and PRT

- The bank uses monies from hiring and initial training to raise the pay for entry-position pay levels.
- Personnel develops a competitive pay package for workers.
- The bank provides workers with advanced training.
- The bank conducts exit interviews to determine reasons for turnover.
- Personnel uses the bank's best workers to train new workers.
- Top management recognizes the difference between turnover and growth.

[Table 25-6](#) illustrates the obstacles and injections that the bank developed for the injection, "Personnel develops a competitive pay package for workers."

The PRT for the injection is illustrated in [Fig. 25-31](#).

A few things to note:

1. It is usually easier to build the PRTs by starting with the most ominous injections (the injections that seem most difficult to achieve). By doing so, you will

typically address the “easier” injections in the process, and you will avoid multiple versions of the same tree.

2. Most of the intermediate objectives and injections are verbalized as entities rather than actions. An objective, whether it is an intermediate objective or a high-level injection, is a condition to be achieved, and an action is something that is done to achieve an objective. The place where we would expect to see IOs written more in the form of actions would be at the “bottom” of the tree; such IOs do not have other IOs pointing to them. At that level, we generally “know what to do,” and the initial obstacles to be overcome are relatively minor. We will see actions in the TRT and as in tactics in the S&T.
3. Each arrow represents the obstacle that exists which is preventing the injection from being achieved. If an IO is pointing to another IO (e.g., 22 pointing to 23), the obstacle (in the arrow that connects them) is also preventing the IO that is pointed to from being achieved.
4. Verify that each obstacle is, in fact, an entity that exists in the current reality of the system. If it does not, it is an imagined obstacle, not a real obstacle, so there is no need to implement an IO to overcome it.

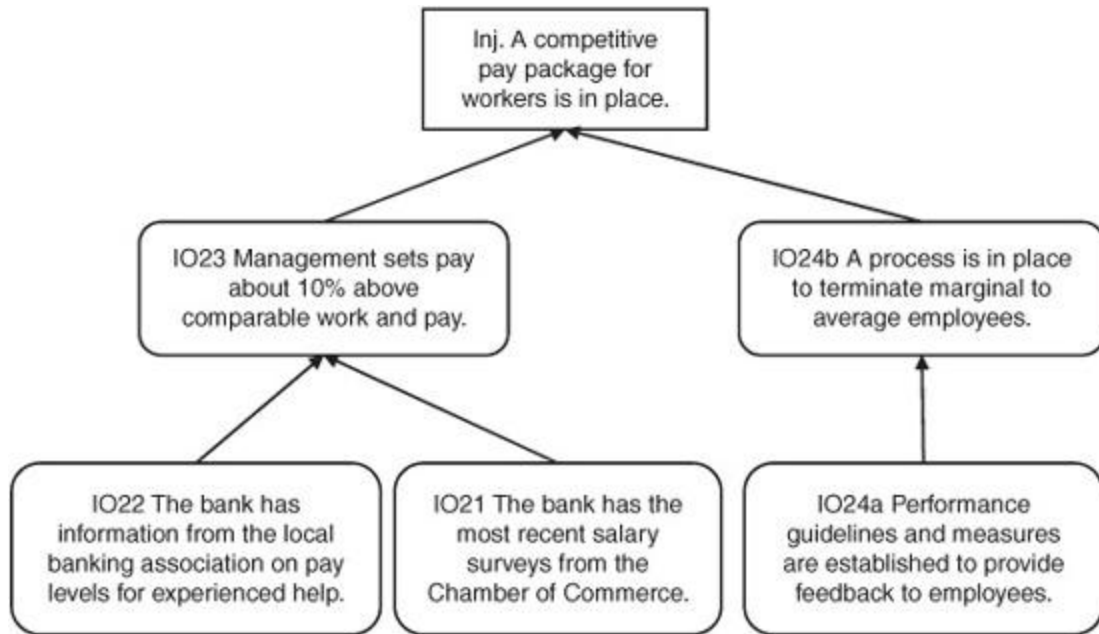


FIGURE 25-31 Bank PRT for injection, “a competitive pay package for workers is in place.”

5. Validate the obstacle causality—is the existence of the entity that is claimed to be the obstacle *really* an obstacle to the achievement of the injection or the IO? If it is not, then there is no reason to implement an IO to overcome it.
6. Verify the IO causality—will the IO really overcome the obstacle and open the door to implementation of the higher IO or injection to which it is pointing? If not, you need to select a different IO.

As the PRTs are developed for each injection, identify any necessary condition relationships that exist among various IOs or injections. This will help you integrate the implementation, rather than simply having a collection of injections to implement. When the bank added to the PRT those IOs it defined to achieve the injection, “The bank conducts exit interviews to determine reasons for turnover,” the PRT expanded as shown in [Fig. 25-32](#).

The full PRT, as the bank team wrote it, is illustrated in Appendix E of this chapter found on the McGraw-Hill website: <http://www.mhprofessional.com/TOCHandbook>.

Transition Tree

Nothing happens until something moves.

—Albert Einstein

We finally reach the place where the rubber meets the road—it's time for action! Some injections and IOs are “no-brainers” to implement. There are others that you know intuitively are risky unless you plan each step in a highly detailed, even choreographed, fashion. For instance, conducting buy-in meetings with other stakeholders in the organization, conducting important sales meetings with buyers or negotiation meetings with suppliers all fall under the category of actions that should be planned meticulously. This is the function of the TRT.

The TRT provides a way to construct an intended action plan (a sequence of actions to be taken) so that the need for each action, the predicted effects of each action, and the appropriate conditions that need to be in place to trigger an action to be taken (and thus the logic of the sequence itself) are all clear. The TRT is useful for planning an important activity, but equally important for monitoring reality during the execution of the plan, so that that we take actions that are needed when they are needed (when the action-appropriate conditions are present), we don't take actions that aren't needed, and we know and are able to pinpoint exactly what and why to modify if reality unfolds differently than the way we had planned. If this seems to be similar to the approach a scientist would take when designing and then executing an experiment, then you have caught on quite nicely!

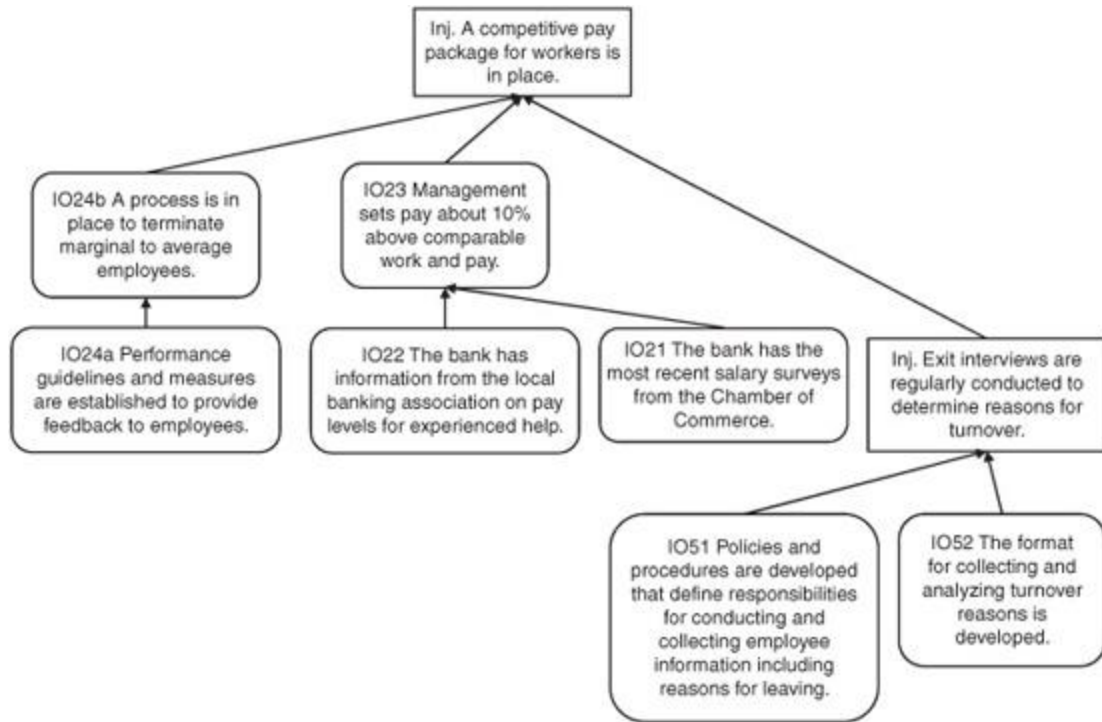


FIGURE 25-32 Bank PRT expanded for injection, “a competitive pay package for workers is in place.”

Never mistake motion for action.

—Ernest Hemingway

The basic structure of a TRT is illustrated in [Fig. 25-33](#).¹⁹ The entities in the tree and the structure of the tree are based on the following concepts:

1. There is a *need* to take an *action*.
2. The fact that an *objective*²⁰ is not yet achieved and will not be reached without additional action means that an action is necessary. In other words, action must be taken because there is some obstacle still blocking the way, and human intervention is required to remove it. By articulating the need for each action, we have an opportunity to check before taking action

to see that the need still exists. (If the need for the action goes away, there is no need to take the action!)

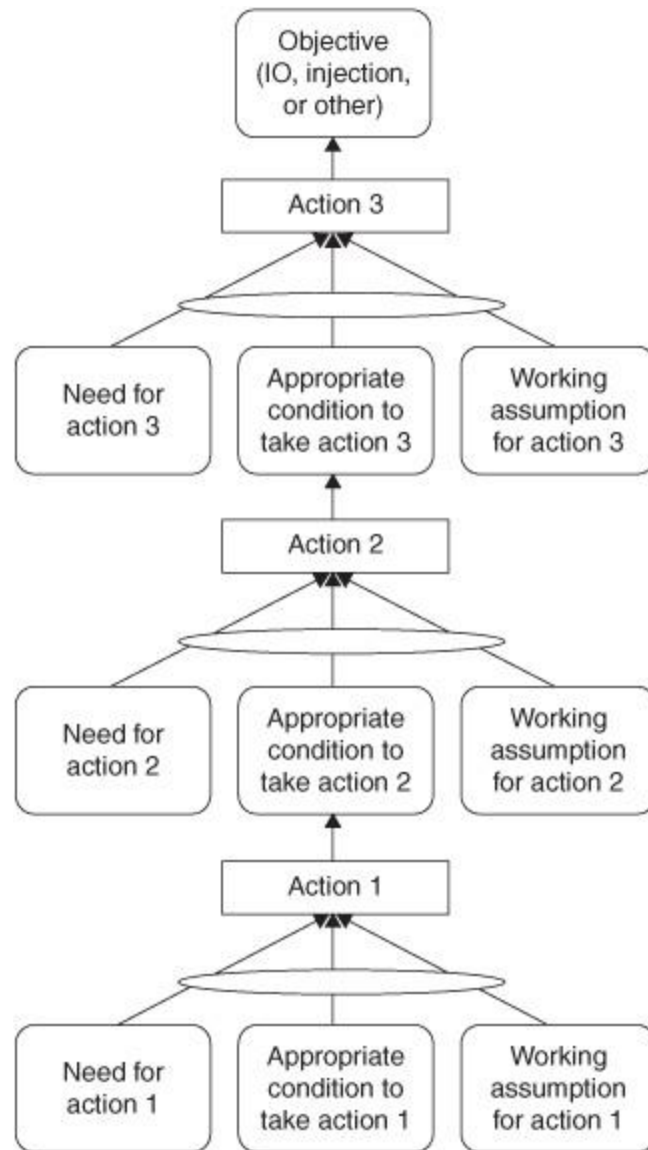


FIGURE 25-33 Basic structure of a TRT.

3. The *conditions are appropriate for taking the action*.
In his July 2001 article, "Transition Tree—A Review," Rami Goldratt articulates what makes *conditions appropriate for taking the next action*.
 - a. *I have the ability to take the next action, and*

b. *The next action will not lead to serious negative effects.*

The sequence of actions is due to the need for the earlier action(s) to cause the appropriate conditions for latter action(s) to be taken.

Let us take a simple example. You are standing at a busy intersection, and the nice restaurant where you are meeting your friend for lunch is across the street. The fact that you are standing on the opposite side of the street from the restaurant means that there is a need for you to take an action, as you must get to the other side of the street. Your first action is to look at the traffic light. The green “OK to cross” signal is illuminated, and traffic has stopped in order to allow pedestrians to cross. The condition is appropriate for you to take your “walk across the street” action, so you confidently do so. On the other hand, if the red “Don’t Cross” signal were flashing, you would know that if you started to walk into the intersection, a car might hit you. In other words, the conditions would have not yet been appropriate, and you would wait a few moments until the light changed in your favor.

The steps to construct a TRT are:

1. Identify the objective and verbalize it as an entity. The objective of a TRT can be an intermediate objective or an injection from a PRT or another objective.
2. Write all of the actions you think should be taken, in the order you expect the actions should be executed, and construct the “spine” of the TRT—the standard protocol is that the first *action to take* is at the bottom of the tree, and the last is at the top. The final action should be pointing to the objective. (See actions 1, 2, and 3, and the objective in [Fig. 25-33](#).) If you cannot

think of any actions, it means that the obstacles are still too big for your intuition to guide you to the actions to take. Go back to the PRT and identify the obstacles and IOs to a lower level—to the point where you have identified an IO that your intuition tells you, “We can do this, and I’ve already got some actions in mind.”

3. For each action, verbalize its associated entity cluster.
 - a. Verbalize the *appropriate conditions* for taking the next action. These are the effects of the action (and are thus the entity to which the action is pointing).
 - i. What negative effects will be caused by the next action, unless I take this action? Verbalize that they will not be created.
 - ii. What new ability do you have after taking the action that brings you closer to the objective and enables you to take the next action? Verbalize the new ability.
 - b. Verbalize the *need* entity.
 - i. What is the need to take this action?
 - ii. Why is this action important? In order to . . .
 - iii. Why take this action? In order to . . .
 - c. Verbalize the *working assumption* entity.
 - i. Why does the *action to take* satisfy the need?
 - ii. What do you assume when you claim that this action satisfies this need?
4. Check the validity of the causality that links each cluster.
 - a. As verbalized, are the *need*, *appropriate conditions*, and *working assumption* that point into an *action to take* sufficient to make the action specified the right action to be taken?
 - b. For any *appropriate conditions* that are intangible or not directly verifiable, identify and map the effects that would be verifiable indicators (“the proof”) that

the *appropriate condition* is in place, as additional effects of the action.

5. Check for negative branches and make the appropriate modifications (modify actions or add new actions in order to prevent the undesired consequences).

In the process of creating a TRT, you may find that you initially indentified actions that really are not necessary. You may also find that you need to add actions that you had not initially thought of in order to close “sufficiency gaps.” You may also find that the sequence you initially had in mind needs some rearranging. How wonderful that you find these things out on paper in the planning stage instead of in reality! Consider how much time and effort you are saving as a result!

I will provide an example of a TRT in the next section of this chapter, to illustrate how a TRT has been used by the sales force of a company that is using TOC to build, capitalize, and sustain a *decisive competitive edge* (DCE).

If anything is certain, it is that change is certain. The world we are planning for today will not exist in this form tomorrow.

—Philip Crosby

The Strategy & Tactic Tree

“The people may be made to follow a path of action, but they may not be made to understand it.”

—Confucius, The Confucian Analects
If an initiative aims to significantly improve an organization’s performance, then inevitably changes would

be needed to various tasks (decisions and actions) that the organization's people are doing. If the initiative is going to stick, then not only the tasks, but the thinking behind those tasks must also change. Irrespective of an individual's level in the organizational hierarchy, or the functional areas in which they reside, each person in the organization wants the same things—to understand how they fit in the big picture, why they are necessary to the whole, and how they contribute to making a real difference.

For each change an initiative requires people to make, they need to understand the changes that they need to make and why. If the answers to the following four questions are not effectively articulated, organized and communicated, people will be forced to make their own assumptions about the answers, and they will behave accordingly. And the likelihood decreases dramatically that the initiative would be a success.

1. For each change I need to make, why do I need to make it?
2. What will the change achieve, vis-à-vis the goal of the initiative?
3. What do I actually need to do in order to make the change?
4. Why will the actions achieve the needed change?

The various TP applications discussed in this chapter provide a robust set of tools with which we are able to fully and logically analyze and describe a core problem, the solution, the hurdles we need to overcome in order to move from the current to the new reality, and even detailed action plans to reach specific milestones and objectives. TOC also provides the recognition of the layers of resistance and an

effective approach to achieving collaboration and buy-in while honoring the win-win principle ([Chapter 20](#)).

But as more TOC implementations focused on holistic organizational transformation rather than single-function improvement programs, it became clear that the standard collection of excellent TOC tools were insufficient to obtain the synchronization and communication required for a major, holistic organizational transformation initiative to achieve and sustain the intended improvements. And they did not provide the means by which anybody in the organization could readily answer the four questions above.

A well written S&T is the TP tool that organizes the full analysis in a way that the answers to the four questions are provided for each function across the organization, to the degree of detail needed at each level up and down the hierarchy, in a single logical map.

The First Step: The Goal

When you look at yourself from a universal standpoint, something inside always reminds or informs you that there are bigger and better things to worry about.

—Albert Einstein, *The World as I See It*.

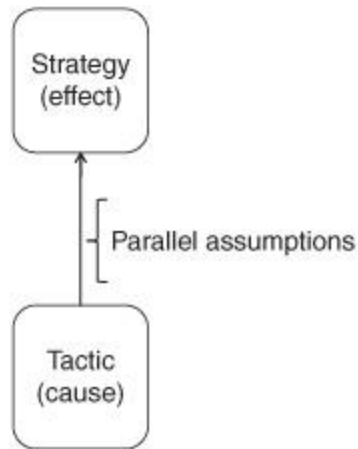


FIGURE 25-34 Cost-and-Effect relationship of Strategy, Tactic, and Parallel Assumption.

Imagine trying to answer any of the four questions for everyone in the organization without first having a clear definition of the goal—the purpose—of the initiative. I can't either. Therefore, defining the goal of the initiative is the starting point of the S&T. For example, the goal of a Viable Vision initiative is stated as follows (with permission from Goldratt Consulting):

The company is an Ever Flourishing Company; continuously and significantly increasing value²¹ to stakeholders—employees, clients and shareholders.

But this high-level statement of the goal does not provide enough information to align and synchronize the specific changes that the organization must make throughout its various levels and functions. We also need a high-level understanding of how the company is going to become ever flourishing. In an S&T, the purpose of the initiative is thus always described with the following three elements:

1. The Strategy—The “What” of the Initiative
 - The purpose of the initiative—the goal the organization is intending to achieve as a result of the implementation.

2. The Parallel Assumptions—The “Why” of the Tactic

- The conditions that exist in reality that lead us to a specific course of action that would achieve the strategy; the logical connection between the tactic and the strategy; a well written set of parallel assumptions explains why the tactic is the course of action that leads to attainment of the strategy.

3. The Tactic—The “How” of the Initiative

- What needs to be done in order for the implementation to achieve the goal.

If you were to model the S&T step using the cause-and-effect mapping process described in this chapter, it would look like [Fig. 25-34](#).

[Table 25-7](#) contains the strategy, parallel assumptions and tactic that comprise the first S&T step for every company that embarks on a Viable Vision implementation:²²

Strategy without tactics is the slowest route to victory. Tactics without strategy is the noise before defeat.

—Sun Tzu

Branching into Layers of Detail

Once the initiative has been defined at the highest level, we can derive the details that are necessary to implement it. Let’s imagine your company is just beginning a Viable Vision initiative, and the CEO has just completed reading to you the strategy, parallel assumptions, and tactic of Step 1 of

the Viable Vision S&T. What is the next set of information that is needed in order to determine the specific tasks that people must carry out to implement the initiative?

Step 1: Viable Vision	
<p>Strategy <i>The "What" of the initiative—the purpose of the initiative; the goal the organization is intending to achieve as a result of the implementation.</i></p>	<p>The company is an Ever Flourishing company; continuously and significantly increasing value* to stakeholders—employees, clients and shareholders.</p> <p>* Increasing value: stability on the green curve, performance on the red curve.</p>
<p>Parallel Assumption <i>The "Why" of the Tactic—the conditions which exist in reality leading us to a specific course of action that would achieve the strategy; forms the logical connection between the tactic and the strategy, explaining why the tactic is the course of action that leads to attainment of the strategy.</i></p>	<ul style="list-style-type: none"> • Realizing a Viable Vision (VV)—reaching results that were considered to be unrealistic while increasing stability; and doing it again—turns a company into an Ever Flourishing company. • For the company to achieve the VV, its Throughput must grow (and continue to grow) much faster than Operating Expense. • Exhausting the company's resources and/or taking too high risks severely endangers the chance of achieving the VV.
<p>Tactic <i>The "How" of the initiative—what needs to be done in order for the implementation to achieve the goal. In a well-written S&T step, the tactic is obvious once the parallel assumptions are read.</i></p>	<p>Build a decisive competitive edge and the capabilities to capitalize on it, on big enough markets without exhausting the company's resources and without taking real risks.</p>
<p>Sufficiency Assumption <i>The "Why" of the next level—explains the need to provide another level of detail to the step; if we don't pay attention to it, the likelihood of taking the right actions is significantly diminished. (explained below)</i></p>	<p>The constraint is management attention. The company must operate based on robust procedures, otherwise the constraint is wasted.</p>

TABLE 25-7 Strategy, Parallel Assumptions, Tactic and Sufficiency Assumptions.

Certainly, the first thing we need is the definition of the company's decisive competitive edge. What is it, and why is it appropriate for your company? What makes it different from the way your company has competed in the past? Once this is understood, the next level of detail must provide the guidance for building it and capitalizing on the decisive competitive edge. Given that this initiative is about

ongoing growth and stability, guidance is also needed on how the company intends to sustain the decisive competitive edge while it grows. For each of these aspects of the initiative, you must then know what it means in terms of the specific changes that you and others must make in your day-to-day jobs, and it is important to assure that the changes you need to make are not in conflict with those above you or below you in the hierarchy, or with other functions.

Notice that your thinking is taking you to increasingly granular levels of detail. Each level of the S&T provides more detail to the level above it. [Figure 25-35](#) illustrates this, and provides the themes of some of the steps that you would find on a typical S&T associated with a Viable Vision implementation.

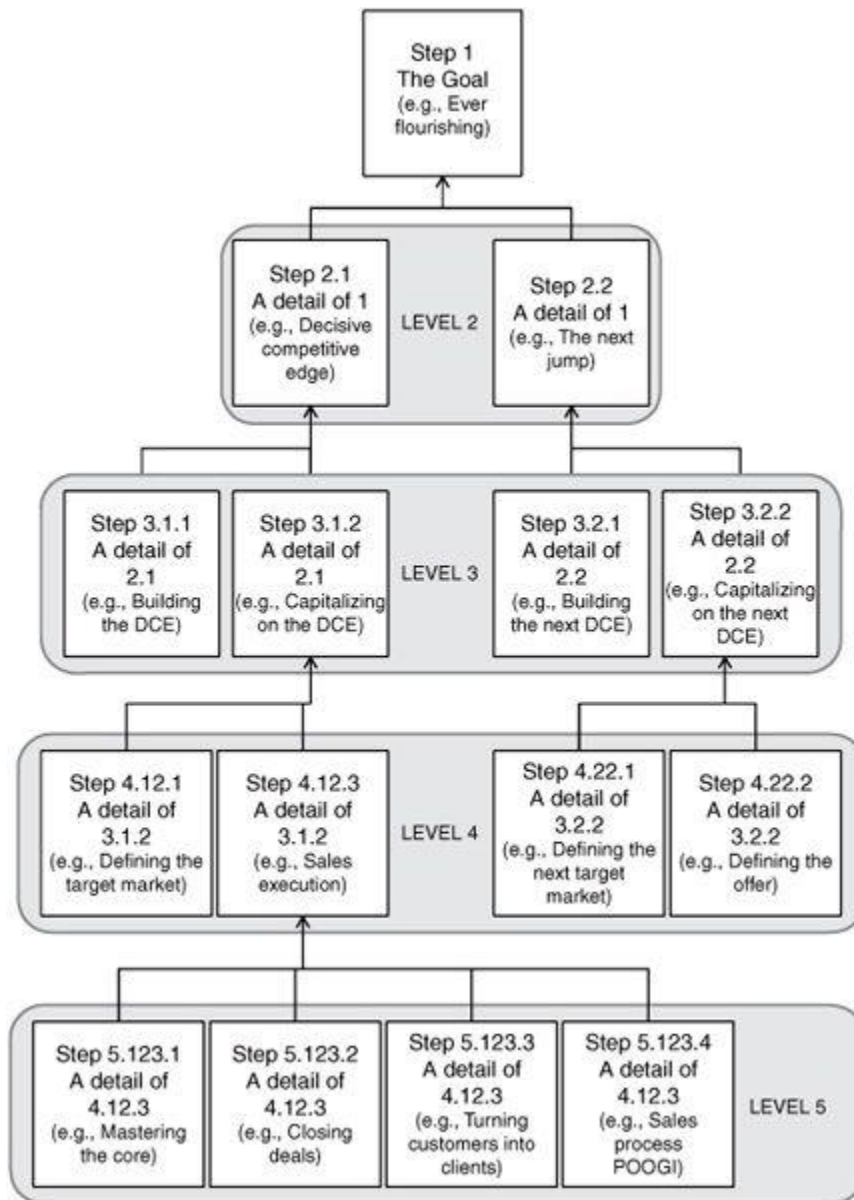


FIGURE 25-35 The S&T cascading levels of detail.

How do we know when a layer should be added? Albert Einstein defined insanity as “doing the same thing over and over again and expecting different results.” Given that what we do is the result of what we think, we can also define insanity as, “thinking the same way over and over again and expecting different results.” The purpose of the initiative is to elevate the organization’s performance. We have already established that this involves making changes not only to

the tasks that people perform, but to the way people think about their tasks and the relationship between what they do and the purpose of the initiative. Therefore, we must consider the potential for inertia—the tendency to think the way we’ve always thought when determining or communicating the changes that must be made to achieve and sustain the intended results of the initiative.

A layer is added only when there is a good chance that inertia will prevent the right actions from being taken. Another way to say this is that if we don’t pay attention to the sufficiency assumption, then the chances of implementing the tactic correctly or achieving the strategy are dramatically reduced. The Sufficiency Assumption is the verbalization of the specific reason for concern. In [Table 25-7](#), you see that the Sufficiency Assumption that guides the next level of the S&T is, “The constraint is management attention. The company must operate based on robust procedures, otherwise the constraint is wasted.”

S&T Elements

Once we have defined of the goal of the initiative as the first S&T step, we have established the single reason for anybody to be asked to make a change to the way they work or think: If they don’t make the change, the organization would be blocked from achieving the goal of the initiative.

As you see in [Fig. 25-38](#), each entity in an S&T is referred to as a Step. From Level 2 downwards, each Step contains several elements:

The Necessary Assumption—The “Why” of the Step

The reason that the higher-level S&T step cannot be implemented unless a change is made.

In other words, it describes the necessity for an action to be taken.

The Strategy—The “What” of the Step

The objective—the intended outcome—of the S&T step. When the strategy is achieved, the need described by the necessary assumption is met.

The Parallel Assumptions—The “Why” of the Tactic

The conditions which exist in reality leading us to a specific course of action that would achieve the strategy; they form the logical connection between the tactic and the strategy, explaining why the tactic is the course of action that leads to attainment of the strategy.

The Tactic—The “How” of the Step

What needs to be done in order to achieve the strategy. In a well-written S&T step, the tactic is obvious once the parallel assumptions are read.

The Sufficiency Assumption²³—The “Why” of the Next Level

Explains the need to provide another level of detail to the step; if we don’t pay attention to it, the likelihood of taking the right actions is significantly diminished.

[Figure 25-36](#) illustrates the necessary and sufficient logical relationships between the various steps in an S&T and their higher and lower levels. In the illustration, both 2.1 and 2.2 are necessary in order for 1 to become reality. Once both 2.1 and 2.2 are implemented, 1 will have been implemented, and the goal of the initiative achieved. Steps 3.11.1, 3.11.2 and 3.11.3 are each necessary for 2.1. Once all three are implemented, the strategy of 2.1 will have been achieved. Steps 3.12.1, 3.12.2, and 3.12.3 are each necessary for 2.2. Once all three are implemented, the strategy of 2.2 will have been achieved.²⁴

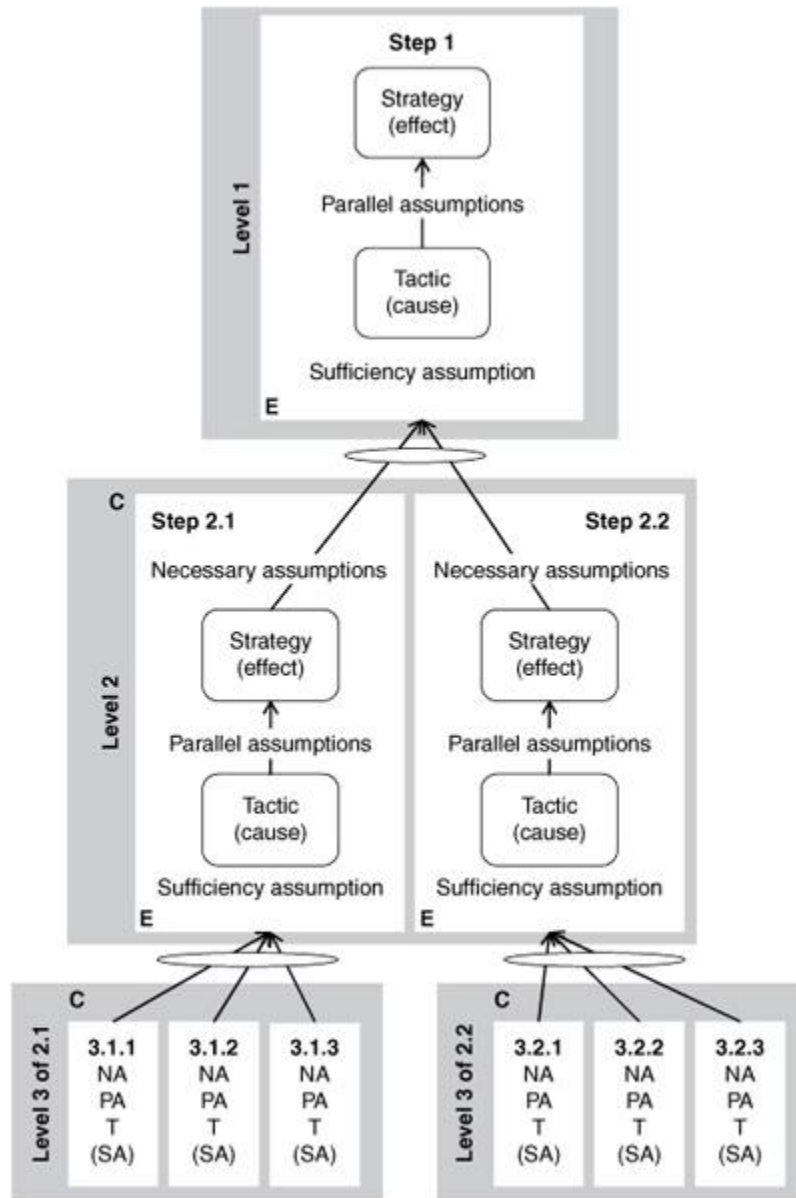


FIGURE 25-36 Logical relationship between steps and levels.

Communication, Alignment, and Synchronization

By using the S&T as the main vehicle to orchestrate and communicate an initiative, the answers to the four questions that people must have in order for an initiative to achieve and sustain its goal are readily available.

1. For each change I need to make, why do I need to make it?
 - This question is answered by the Necessary Assumption.

2. What will the change achieve, vis-à-vis the goal of the initiative?
 - This question is answered by the Strategy.

3. What do I actually need to do in order to make the change?
 - This question is answered by the Tactic.

4. Why will the actions achieve the needed change?
 - This question is answered by the Parallel Assumptions.

By examining a branch of the tree vertically, we see the alignment of each level in the hierarchy. By examining a S&T horizontally, we see the synchronization across functions.

The structure of the S&T provides the way for us to understand how any local action is contributing to the global goal of the initiative.

Implementing an S&T

People love chopping wood. In this activity one immediately sees results.

—Albert Einstein

Just as with the rest of the TP and TOC, the logic of the scientist is applicable to the use of the S&T. If an assumption in the S&T is found to be invalid in the environment in which the S&T is being implemented, then it is likely the corresponding strategy or tactic should be changed! Therefore, it is crucial to ensure that from the beginning of any implementation, the assumptions are being checked and validated, and that as actions are taken, the intended effects are checked.

The activities of an implementation for any higher level step in the S&T are defined in the lowest level that has been written for the step. The S&T is written so that the order in which the activities take place can and should generally be implemented from left to right. One of the most important rules governing the best in class implementations of S&T's is "one step at a time." Following this guideline provides for the ability to:

- Check that the cause and effect assumed in an S&T step is what actually occurs in the reality of the implementation. Remembering the cause-and-effect relationship between the tactic and its strategy, we know that once we implement a tactic, we should be able to verify that the strategy—the objective of the tactic—is in place. There are only a few reasons why it would not be:
 - The tactic was not implemented correctly.
 - There is another aspect of reality that was not taken into account, which is blocking the strategy from being in place.
 - The parallel assumption was incorrect.
 - Implementing one step at a time makes it exponentially easier to check for each of these possibilities and make the appropriate course corrections very quickly, and with clear understanding. For each additional step we

- try to implement simultaneously, the number of variables we must check increases significantly, our chances for incorrect assessment of the problem increases, and the time that must be spent on analyzing, checking, and correcting increases.
- Secures the understanding of the cause-and-effect relationship that exists between the tactic and strategy. It is one thing to read an S&T or to get instructions and training to implement a specific tactic. It is quite another to actually experience the positive effects from implementation of a specific tactic. When it is crystal clear that a specific action or set of actions leads to a specific significant improvement for those involved in the implementation, the inclination to “go back to the old way” is increasingly blocked.
 - Helps to avoid bad multitasking, which always leads to increased timelines and mistakes.

Combining steps lengthens the time to secure the results and puts the implementation at risk. Taking one step at a time helps to ensure that the good changes will stick, and the not-so-good changes (surprises) can be addressed immediately because the cause is known.

Appendix G contains screenshots of the hierarchy of an S&T used for many Make-To-Order companies.²⁵ The activities under 3.1 generally fall under Operations, and the activities of 3.2 generally fall under Sales and Marketing, so the implementation of 3.11 and 3.12 can occur simultaneously. Under 3.1, the implementation goes in the order of the Level 4 entities—4.11.1 through 4.11.6. Some of these Level 4 entities are detailed to Level 5 and some are not. Entity 4.11.1 is implemented at Level 5, starting with 5.111.1, and completing with 5.111.4. Only after 5.111.4 is completed, and the strategy of 4.11 is verified to be in place, do we move to 4.11.2 through its Level 5

entities, 5.112.1, 5.112.2, and 5.112.3. After checking that the strategy of 4.11.2 is in place, we move to 4.13, etc. The same approach is used to implement 3.12. We begin with 4.12.1, via its Level 5 entities 5.121.1 through 5.121.4.

We cannot do everything at once, but we can do something at once.

—Calvin Coolidge

Using the TPs to Implement an S&T

Earlier in the chapter I showed you one example of how the Cloud had been used in implementing the POOGI step of an MTO S&T. I would like to provide a few more examples of the use of the TP when an S&T is guiding an implementation.

Use of the Negative Branch Reservation

As you can imagine, implementing a major initiative requires gaining and sustaining the understanding of and buy-in to what is being implemented, and in the cases of valid reservations, making the appropriate modifications to the S&T. The NBR is used to facilitate the accomplishment of this, both before and during the implementation. The example I provide here is taken from a manufacturing company in the United States that made the decision to go on a Viable Vision implementation. The critical part of the decision process is for the top management and other key people in the company to go through a multi-day session in which they learn the relevant TOC and scrutinize their S&T. At key points in the session, they map out the NBRs that are concerning them about implementing specific aspects of the S&T. The company's S&T is a modified MTO S&T, and one of the reservations that was expressed during the session was focused on the tactic of Step 3.1, which states:

The company manages its operations according to the four concepts of flow.

The NBR that the manager submitted is pictured in [Fig. 25-37](#). When they looked at the NBR they had written, it became obvious to the management team that they were the key to preventing the negative effects from emerging. Their injections, which were incorporated in their implementation plan, were to ensure that education was provided to employees and management alike, and the commitment that if any existing measures turned out to reinforce the belief, they would be addressed. They were relieved that they could deal with the issue, and were energized to continue the session and move into their implementation. Eighteen NBRs were documented and addressed by the management team in that session. Some needed injections, two resulted in modifications being made to the S&T, and most were addressed by gaining better understanding of the S&T, itself.

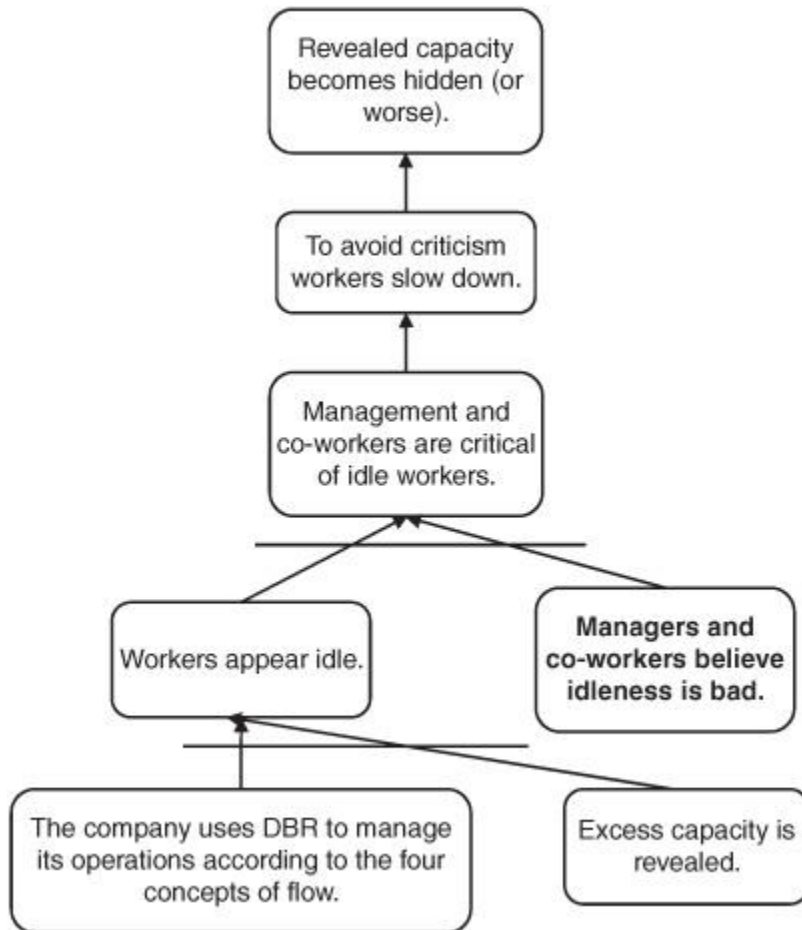


FIGURE 25-37 NBR for “managers and co-workers believe idleness is bad.”

Use of the Transition Tree

When a company has a DCE, it is solving a significant need of its market to the degree that none of its major competitors can. A customer need that is not addressed by any of the significant suppliers in a given market is not something that the suppliers tend to emphasize in the sales process. It is also not something the customers in the market emphasize, given that the suppliers do not address it. This means that in order for a company to really capitalize on a DCE, it must make some fundamental

changes to its sales process—changes that will highlight the need and the company’s unique ability to address it.

The S&T for a company whose DCE is reliability (of due dates) provides the instructions for a core meeting between the salespeople and potential clients who would appreciate the company’s offer of reliability. The TRT is used to design (choreograph) that meeting. Step 5.123.2 in one such tree is provided in [Table 25-8](#) (See Appendix F):

Step 5.123.2 Mastering the Core	
Necessary Assumption	<ul style="list-style-type: none"> • Achieving the client’s strong buy-in to the great value of the offer is the core of reliability selling (performing it properly boosts the sales process, performing it poorly almost guarantees failure). • The client has a set expectation of what the vendor is supposed to present in the first sales meeting. Following the set expectation of the client and just presenting the offer (without the supporting logic) guarantees failure.
Strategy	<ul style="list-style-type: none"> • Salespeople are skilled at conducting the raising interest meeting—the core of reliability selling—getting the buy-in on the great value of the offer.
Parallel Assumptions	<ul style="list-style-type: none"> • Vast experience shows that raising interest meetings are successful if constructed along the following lines: <ul style="list-style-type: none"> • The value of the reliability offer is in eliminating problems—the damage caused by delays. Getting a consensus that meaningful damage of delays exists is the first key step in obtaining the buy-in. Presenting the damage as a result of common practices in the supplier’s industry strengthens the perception of the company as a reliable supplier looking to bring value to its clients. It also prevents the risk that the client will argue the existence of the problems to avoid admitting failures in his or her area of responsibility or to avoid giving power to the supplier in the “negotiation game.” • Presenting a list of sensible criteria to judge any suggested solution, aiming to eliminate the damage, is an effective technique to pave

	<p>the way for the client to recognize the reliability offer as the obvious best solution to his or her problem. It also blocks any unsatisfactory different directions for a solution that the client may entertain.</p> <ul style="list-style-type: none"> • The bitter experience with unreliable suppliers conditioned clients to look for “the snake in the grass”—examining carefully the offer elements, checking if it solves the problems, if it does not involve real risks, and if it is practical to implement. An effective way to strengthen the position of the company as a reliable supplier is to unfold the offer elements as best meeting the criteria. • Using the client’s remaining concerns (spoken or unspoken) as the base for the next steps (in which the concerns will be decisively put to rest) contributes significantly to the reliability perception. • Role playing is an effective technique to master a new buy-in process: “The more you sweat the less you bleed—difficult in preparation, easy in battle.” • The most effective way to convince the sales force that such a radical sales meeting does work is to cause the team to experience it firsthand.
Tactic	<ul style="list-style-type: none"> • The reliability core meeting is designed by key salespeople. • Key salespeople are coached (extensive role-play) and handheld until they personally achieve successful core meetings—the pre-launch. • Note: If the green light is not given yet, the offer is presented as a future service the company is about to launch. (The company can even establish deals with a future activation date.)
Sufficiency Assumption	None.

TABLE 25-8 Step 5.123.2 Mastering the Core.

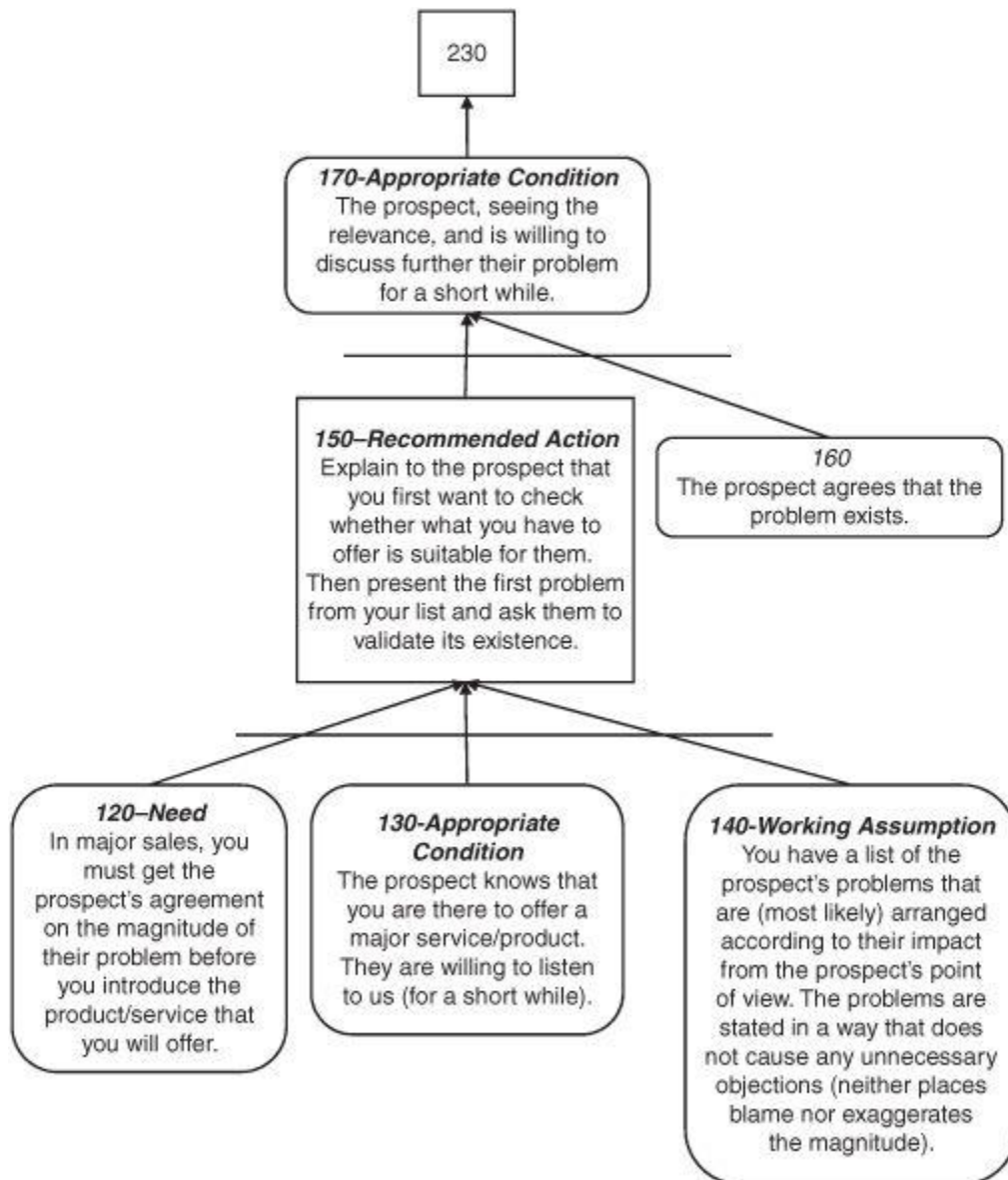


FIGURE 25-38 TRT cluster from reliability selling example.

I am providing a section of the TRT that was developed for the salespeople to use in order to learn how to conduct the meeting, and to debrief each meeting that they held.²⁶ It is based on the specific verbiage contained in the S&T step above, along with the TOC Expert's knowledge of the Layers of Buy-In,²⁷ and thus to prevent any objections that would otherwise be raised (see [Fig. 25-38](#)), such as:

1. You don't understand what the problem is.
2. I/we don't agree on the direction of the solution.
3. Your solution can't possibly deliver the level of success you claim (too good to be true).
4. Your offer will cause bad side effects.
5. Even if I/we wanted to do this, there are obstacles that block us from implementing the solution (actually buying from you).
6. Other un verbalized fears.

Action will remove the doubts that theory cannot solve.

—Tehyi Hsieh

From TP Analysis to S&T

Every assumption in the S&T should be an entity that is part of the current reality, and can (should) be validated as such. Therefore, the assumptions can be found in the CRT, the EC, and the obstacles of a PRT. The FRT provides the strategy at the highest level, which is essentially the summation of the *desired effects* (DE), and NBRs provide the input to Level 4 of the S&T. Level 5 comes directly from the obstacles that are verbalized in the PRT process and thus frames at the lowest level the initial actions to be taken to achieve the strategy. [Table 25-9](#) provides a cross reference between where you will often see elements of the TP analysis and the components of S&T steps. The S&T will present the various elements in the form of the actual entities, causalities, and summaries of the various trees or branches of trees. While you will see elements of the PRT and TRT in the S&T, it is

typically not necessary to create complete PRTs or TRTs in the process of creating an S&T.

	Necessary Assumption	Strategy	Parallel Assumptions	Tactic	Sufficiency Assumption
Current Reality Tree	UDE's		Entities and causalities that exist in current reality		Causalities
Evaporating Cloud	Conflict (D-D')	A, B, C entities Injections		Injections	Assumptions
Future Reality Tree		Summary of desired effects Injections	Entities and causalities that exist in current reality	Injections	
Negative Branch Reservation		Injections	Entities and causalities that exist in current reality	Injections	
Prerequisite Tree (replaced by S&T)	Obstacles	IOs		IOs	
Transition Tree (replaced by S&T)	Need entities		Working assumptions	Actions	

TABLE 25-9 Cross Reference Between the TPs and the S&T tree.

The Knowledge Organizer

I hope that I have conveyed how a well-written S&T can provide an organization with the ability to achieve levels of communication, synchronization, and performance not previously thought possible. It organizes the answers to the three questions of change in a single document, providing cascading levels of logic and detail needed by each level and function in the organization. By making every

assumption explicit, it provides a means by which we can exercise the mentality of the scientist and carry out our implementations with confidence. Personally, I can no longer envision leading or participating in a major change effort without using the S&T as the blueprint and roadmap for the initiative.

We are rapidly learning more and more applications for the S&T. For instance, by the time the next TOC Handbook is published, we should be able to provide the detailed guidelines for using the S&T to analyze and define an organizational structure, and to analyze and detail the scope of a project. Stay tuned!

Chapter Wrap-Up

Dr. John Grinnell's Project Leadership Model (2007), depicted in [Fig. 25-39](#), is an appropriate aid to conclude this chapter on the TOC TP.

Every organization has a goal. Achievement of a goal is an effect—a result—of actions taken by people.

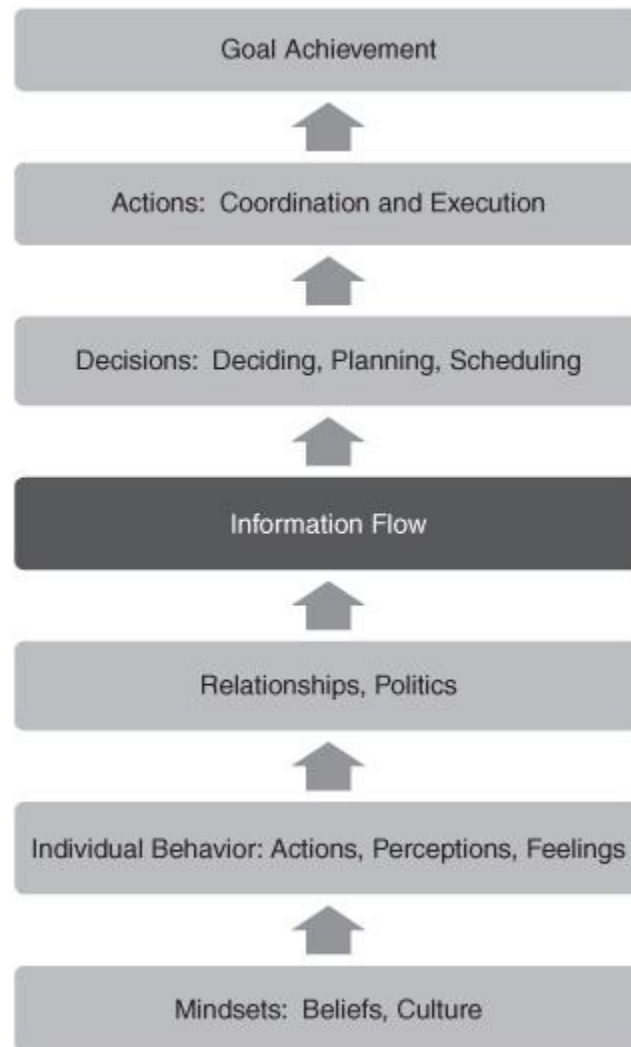


FIGURE 25-39 Project leadership model.

The actions that people take are also effects—results—of the decisions that people made to take the actions.

Decisions are made based on the information available to the persons making the decisions. The point at which information flows is what Grinnell (2007) refers to as the “pinch point” because it comes at the transition between the tangible stuff that we measure, manage, and engineer, and the personal stuff that nobody sees. Let’s dive down. Grinnell’s claim, which, frankly, I can’t argue with, is that the clarity and availability of information is a function of the relationships among the deliverers and receivers of the

information. The quality of the relationships between deliverers and receivers of information is an effect of their perceptions of one another, and the foundation for those perceptions is mindsets that stem from the beliefs and culture of the individual.

The use of TOC tends to be focused on those things that are above the information line in Grinnell's model. However, the actual use of it—which means starting with the concept of inherent simplicity and the mindset of a scientist, the acceptance of the possibility that people are good, and the discipline of internal honesty—has a tremendous impact on those things below the “information line.” The use of TOC TP will change your feelings, behaviors, and relationships, and the result is greater harmony.

References

- Cox III, J. F., Blackstone Jr., J. H., and Schleier, Jr., J. G. 2003. *Managing Operations: A Focus on Excellence*. Great Barrington, MA: North River Press.
- Goldratt Consulting Ltd. 2009. POOGI for MTO Manufacturers, MTO S&T.
- Goldratt, E. M. 1990. *What is this thing called Theory of Constraints and how should it be implemented?* Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 1994. *It's Not Luck*. Great Barrington, MA: North River Press.
- Goldratt, E. M. 2009. *The Choice*. Great Barrington, MA: North River Press

Goldratt, E. M. and Cox, J. 1986. *The Goal*. Rev. ed. Croton-on-Hudson, NY: North River Press.

Goldratt, R. 2001. "Transition tree—A review," (unpublished) Kfar Saba, Israel.

Grinnell, J. R. 2007. *Project Leadership Model*. Chapel Hill, NC: Grinnell Leadership & Organizational Development.

Newton, I. 1729. *The Mathematical Principals of Natural Philosophy*. Volume II. Translated to English by Andrew Motte.

Scheinkopf, L. J. 1999. *Thinking for a Change: Putting the TOC Thinking Processes to Use*. Boca Raton, FL. St. Lucie Press.

conflict.Dictionary.com. *The American Heritage® Dictionary of the English Language, Fourth Edition*. Houghton Mifflin, 2004. <http://dictionary.reference.com/browse/conflict> (accessed: December 19, 2009).

situation.Dictionary.com. *The American Heritage® Dictionary of the English Language, Fourth Edition*. Houghton Mifflin Company, 2004. <http://dictionary.reference.com/browse/situation> (accessed: December 18, 2009).

system.Dictionary.com. *The American Heritage® Dictionary of the English Language, Fourth Edition*. Houghton Mifflin Company, 2004. <http://dictionary.reference.com/browse/system> (accessed: December 18, 2009).

About the Author

Lisa Scheinkopf is a Director of Goldratt Consulting, and is recognized worldwide as a leading Theory of Constraints (TOC) authority. Lisa worked with Dr. Eliyahu Goldratt in developing the TOC Thinking Processes and is the author of the definitive TOC reference, *Thinking for a Change: Putting the TOC Thinking Processes to Use* (St. Lucie Press, 1999). Her articles have been published in a variety of industry and professional publications, and she has a long history of implementing, teaching, and public speaking on TOC. With over 25 years management and consulting experience, Lisa is a past Board Member and Chairperson of TOCICO, and has an MBA in International Management from the Thunderbird School of Global Management.

Appendix B: Categories of Legitimate Reservation

The Categories of Legitimate Reservations—The Rules of Logic[28](#)

Goldratt developed a set of logic rules, called the *categories of legitimate reservations* (CLR), to improve communications when using the TP. The purposes of the CLR are to check your logic in constructing your own diagrams and to check the logic of another person's diagrams. They provide a precise methodology for pinpointing errors in your or another person's thinking. The CLR relate to entities or statements in a logic diagram. Three levels of categories of reservations exist. Each level probes deeper into investigating the logic structure. Many of these concepts are difficult to understand at first, but with a little practice, they become second nature. We provide the three levels and seven categories of reservations with examples in Fig. 25-B1. We will revisit these reservations again in this chapter

as we present and illustrate each tool. Read each example provided in Fig. 25-B1.

Level 1 Reservation (Clarity)

Clarity is used to develop a better understanding of an entity (a logical statement), the causality between two entities, or an area of the diagram. In studying a diagram and encountering any problem, the *clarity reservation* is used. It is always the first reservation used. You are asking the presenter to clarify so you can understand better (the cause entity, the effect entity, the causality connecting the two, an area of the diagram, and so on). For example, in Fig. 25-B1, the reviewer may not understand an entity such as 10 or 20, or she may not understand the causal linkage between 20 and 10, or she may not understand a whole segment of the diagram such as 20, 30, and 10. The reviewer would ask for *clarity*. If the presenter's explanation is unsatisfactory, then the reviewer should use one of the Level 2 reservations to pinpoint the misunderstanding.

Level 2 Reservations (Entity Existence and Causality Existence)

The entity existence and causality existence reservations are used to determine if the entity or statement itself exists or if the causality relationship exists. Examples are provided in Fig. 25-B2.

Entity existence reservation is challenging the existence in reality of either the cause entity or the effect entity. For example, entity 25 is an incomplete sentence. In that state, it is difficult to determine if the entity exists at all. In addition, the reviewer could challenge whether an entity exists in the current environment—*entity existence reservation* for entity 10. The reviewer does not think that entity 10 “Competition is fierce for our product” exists. She

offers as evidence that our company has higher quality and lower prices than competitors do.

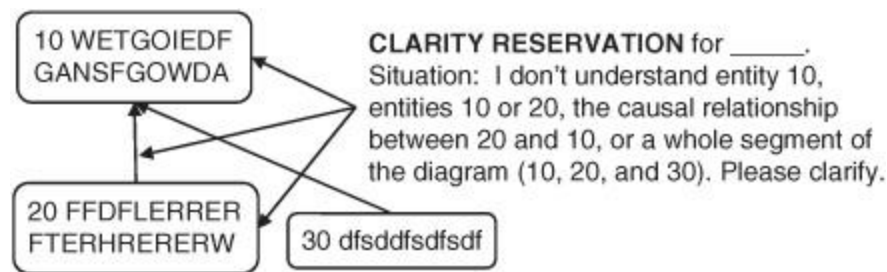


FIGURE 25-B1 Level 1 Reservation (Clarity).

Causality existence reservation is challenging whether causality exists between the two entities. It is challenging the causal arrow—Does the cause entity really cause the effect entity? The example in Fig. 25-B3 provides a situation where the reviewer does not believe that entity 10 “Competition is fierce for our product” is the cause of entity 25 “Our firm is experiencing low profits.”

If the presenter’s explanation is unsatisfactory in showing the existence, then the reviewer should use the Level 3 reservations to pinpoint the misunderstanding. At Level 3, the reviewer must be ready to challenge the logical relationship using a specific reservation.

Level 3 Reservations (Additional Cause Reservation, Cause Insufficiency Reservation, House on Fire Reservation, and Predicted Effect Existence Reservation)

Level 3 challenges should only be used after applying the previous two levels.

The *additional cause reservation* is used to challenge that the presenter has captured the major causes of the effect entity. It is begging the question that there is at least another cause that creates at least as much damage as the current cause entity. A “magnitudinal and” connector is utilized to satisfy this reservation. Each cause entity independently contributes to the effect entity. If cause entity

then effect entity. If (additional) cause entity then effect entity. This situation is indicated where two or more arrows enter an entity and have no “and” connector. Each cause independently contributes to the effect’s existence. In this situation, all causes must be eliminated to eliminate the effect. In Fig. 25-B4, the reviewer believes that 15 “Material costs have doubled in the last quarter” has at least as significant an impact on 25 “Our firm is experiencing low profits” as does the suggested cause of 10 “Competition is fierce for our product.”

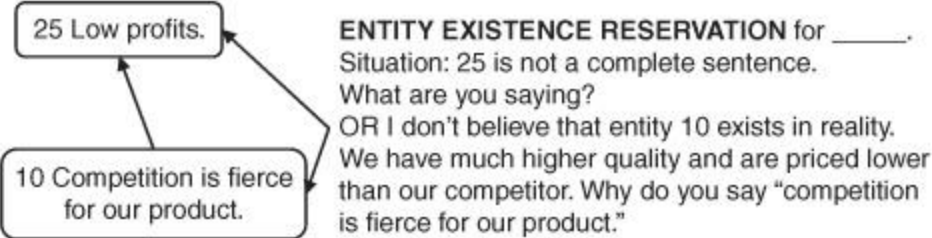


FIGURE 25-B2 Entity Existence Reservation

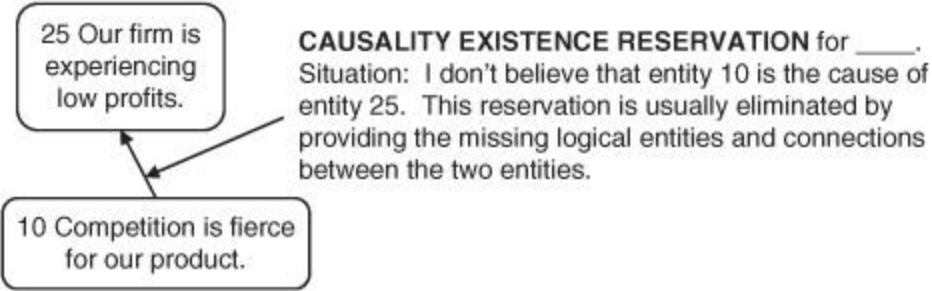


FIGURE 25-B3 Causality Existence Reservation

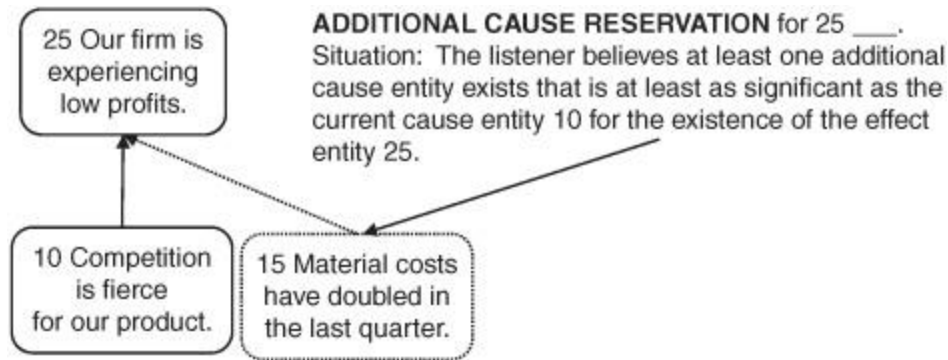


FIGURE 25-B4 Additional Cause Reservation

By using the *cause insufficiency reservation*, the listener is indicating that he or she believes that the current cause entity is insufficient by itself to cause the effect entity. It is begging the question that something else must also exist in addition to the current cause to create the effect. A “conceptual and” connector is usually required to satisfy this reservation. If cause entity and entity (or core driver) then effect entity. The connector is diagrammed as an ellipsis (sometimes called a banana) or line across the arrows. In Fig. 25-B5, the reviewer is challenging that entity 10 “We have not settled on a new union contract” could cause 25 “Our employee morale is low.” She suggests that a more accurate explanation is: If 15 “The current contract expires at the end of the month” and 10 “We have not settled on a new union contract” then 25 “Our employee morale is low.”

The *house on fire reservation* (sometimes called the cause-effect reversal) is used to challenge the thought pattern where the cause and effect seem reversed. This usually occurs where the presenter confuses why the effect entity exists with how we know that the effect entity exists. For example (see Fig, 25-B6), if (cause) smoke is billowing from a house then (effect) the house is on fire is not valid logic. An electrical short circuit may cause the house being on fire. If (cause) the house wiring had an electrical short

circuit then (effect) the house is on fire. The cause of the fire is a short circuit in the electrical wiring. The original statement is how we know the house is on fire, not the cause of the fire. The smoke billowing from the house is the result of the house being on fire. We have confused the cause with the effect. Ask “why” to determine the cause.

The *predicted effect existence reservation* is used to explain why you disagree with the presenter’s previous explanation and generally is the last reservation used. In this challenge, you are prepared to show the presenter that his or her logic is flawed. There are two types of challenges—one questioning the existence of the cause entity and the other questioning the existence of the causality between the two entities. This challenge is presented by providing a counter example that if the predicted effect is present, then the cause cannot be present or if the effect is absent, then the cause cannot be present. In Fig. 25-B7, If 10 “Our quality has deteriorated significantly” then 25 “Our profits have decreased significantly” would be validated by the existence of 35 “Our returns and field service expenses have increased significantly.” However, in examining our expenses this effect does not exist. The reviewer then challenges the existence of entity 10. Suppose the cause entity exists—what other predicted effect must be present? If that predicted effect is not present, then the cause is not present. Likewise, if the predicted effect exists it adds validity to entity 10 being the true cause of 25.

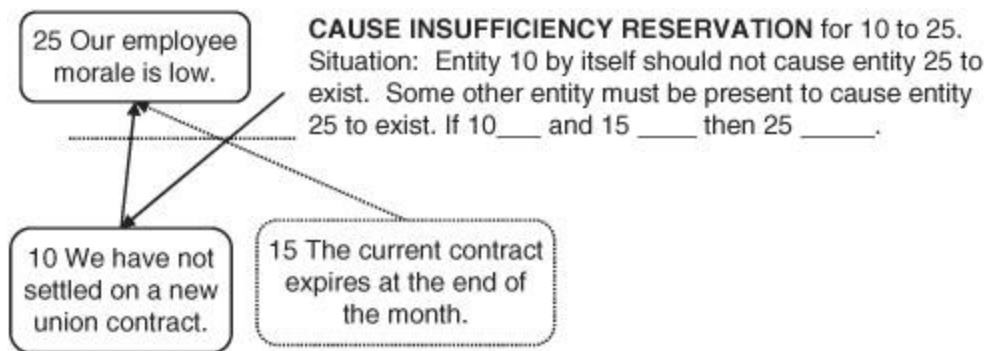


FIGURE 25-B5 Cause Insufficiency Reservation

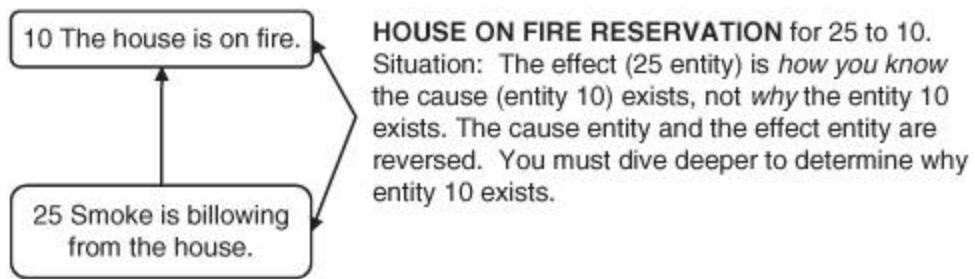


FIGURE 25-B6 House on Fire Reservation

The challenge can be based on the existence of the causality—*predicted effect reservation* for 10 to 20. In the example in Figure 25-B7, If 10 The packaging line broke down then 20 The AJAX shipment is late is challenged for causality—while the reviewer believes that both 10 and 20 exist, she does not believe that 10 caused 20. She offers as proof that the packaging line broke down after the AJAX order was completed; therefore, the line breaking down did not cause the order to be late.

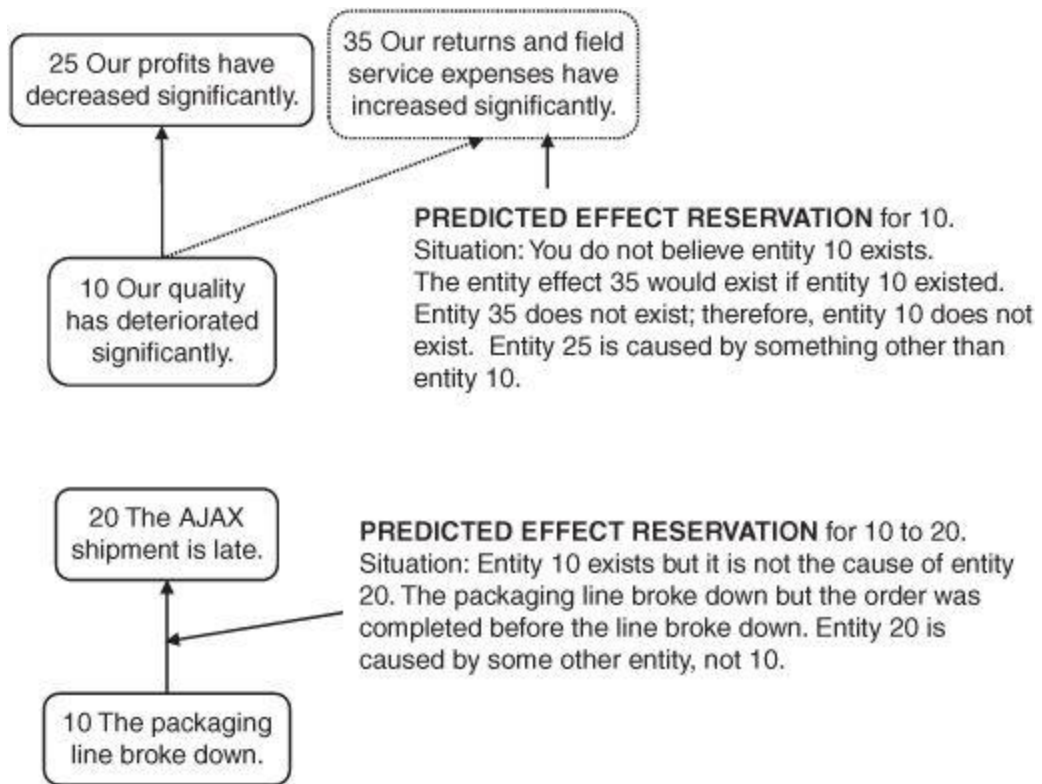


FIGURE 25-B7 Predicted Effect Reservations

CHAPTER 26

TOC for Education

“. . . To Make the Wish Come True”

Kathy Suerken

Why Change?

“When and why did you decide that these thinking tools would work with children all over the world?” asked the Mexican educator through a Spanish translator at the 2001 Mexican TOC for Education Conference. Mrs. Gonzales and 300 other stakeholders of the Nuevo Leon school system had just witnessed very convincing evidence of the efficacy of the TOC tools to enable students of all ages and skill levels to take responsibility for their own learning and behaviors. Moreover, not only were children and educators applying these problem-solving tools outside the classroom to improve family relationships but also some educators and especially those involved in supportive social services were finding these tools very effective in resolving situations of child abuse and to rehabilitate students in juvenile justice institutions.¹

Thus, although Mrs. Gonzalez’s question had to be translated for me, the reasons for it did not. Most people are naturally curious about the origins of a program that brings such broad and deep positive change—especially one that works with so many diverse students and adults. The problem of how to differentiate instruction to students with disparate levels of knowledge, experiences, and interests

within existing resources is the one dilemma most commonly cited by teachers when asked during *TOC for Education* (TOCfE) seminars and workshops on five continents.

So what was the compelling evidence that convinced *me* of the potential global impact of TOC for children? Did I begin to realize the power of TOC as a teaching methodology when I observed the effect of these powerful thinking tools with my own mainstreamed² middle school students, including those considered to have learning disabilities and other special needs? Was it when I realized other local educators were getting similar results with a variety of age groups and even in interventions with very disruptive students? Or was it when it came to my attention that not only were these students teaching these thinking tools to their peers but even to—and at the request of—their parents?

There is a common denominator for these successes—one that is not dependent on unique teachers or circumstances but rather on a methodology demonstrated in the book *The Goal* (Goldratt, 1984). Although many consider it to be a business novel about production, as a teacher, I found *The Goal* to be a book about education—learning to learn, learning to think, learning to lead. I was captivated by the methodology used to enable others to think for themselves, solve their own problems, and take ownership of implementing solutions. While this methodology is not new, what was new to me was the way the scientific method and Socratic questioning techniques were used to motivate others to be more productive and responsible for outcomes in their everyday lives.

After writing the author, Dr. Eli Goldratt, a thank-you letter to explain how I had begun to use this approach to education within my social studies classes and, as well, in managing a volunteer schoolwide international math

project, I received, on behalf of my students, a scholarship to a formal training of the TOC thinking processes taught through applications to business and industry. A facilitator training soon followed to enable me to share this knowledge with other local educators.

Later, when teaching 7th grade students a pilot TOC critical thinking class, I shared how grateful I was for this opportunity, along with my concern that I could never repay Dr. Goldratt and the Avraham Y. Goldratt Institute for this expensive, invaluable learning experience. The students suggested an alternative way to express my gratitude . . . a payment in kind. Thirteen-year-old Jesse Hansen converted an idea into a viable solution with words that succinctly and profoundly convey just how much children, like those who teach them, want to make a meaningful difference. “You can use us, Mrs. Suerken. You can use our work.” Their work became a full range of powerful examples of the tools and their impact and was shared by these students, along with the work of several local educators, at a 1994 TOC business conference attended by Dr. Goldratt.³

In taking note of how effectively the TOC thinking processes could be translated into practical and highly beneficial outcomes in a classroom and, in keeping with his own lifetime goal, Dr. Goldratt created TOC for Education (TOCfE) in 1995 as a not-for-profit organization to disseminate the TOC logic-based tools and common sense methodologies to all who educate others. Since then, TOCfE has reached more than 200,000 adult education stakeholders with an impact on more than 8 million children in 21 countries.⁴ Just like the explanation needed to reveal why these tools work with children all over the world, perhaps the most important ingredient in how TOCfE has continued to grow, develop, and continuously improve worldwide is not so much the timeline but the *why-line*.

In TOC, the whys of creating change that leads to desired and ongoing improvements requires the examination of three questions:

What to Change?

What to Change to?

How to Cause the Change?

The purpose of this chapter is to apply these three questions to the education of children and to answer them by using Goldratt's *Thinking Processes* (TP). This framework will also provide the organization of the chapter, which concludes with a summary.

What to Change?

Many times, we create solutions for problems without first really understanding what causes them. In such cases, we may end up with temporary or partial fixes and the problems resurface. Thus, there is an important distinction between solutions that bring change and solutions that bring improvements. As Eli Goldratt describes this reality, "Although every improvement is a change, not every change is an improvement."⁵

Most educators can recount a litany of solutions and reform programs that have brought considerable change to schools but not the envisioned improvements needed to prepare all children sufficiently to become productive and responsible adults. Thus, in spite of all best practices and the good intentions and hard work of talented, dedicated educators, many symptoms of an elusive core problem remain, such as:

- Many students do not know how to connect, interpret, and question information in what they read or hear.

- Many students memorize rather than analyze information.
- Many students do not know how to solve problems and are dependent on others to do so for them.
- Some students do not perceive what they are learning to be relevant to their lives and therefore disengage.
- Many students do not know how to apply what they learn.
- Many students do not think through consequences before taking actions.
- Some students do not know how to control impulsive behaviors that sometimes lead to violence.
- Some students leave before graduating.
- Maintaining the highest standards for meeting the learning and behavior needs of all students requires more resources (especially time) than are currently available to educators.

These ongoing undesirable effects impact all education stakeholders who look to the education system to prepare youths to be responsible citizens and productive workers in an increasingly competitive, global marketplace. Therefore, with so much at stake, when changes do not lead to desired and expected improvements, there is understandable disappointment and frustration. Unfortunately, these outcomes usually result in explanations written in the language of blame, which are typically directed at those considered being responsible for implementing the chosen solutions, even if they were not part of the process of creating them. If we can—and should—assume that those in education want to be good educators, then it is also reasonable to assume that they are justifiably sensitive to criticism that impugns their abilities, motivation, and especially their purpose.

Educators feel overwhelmed by the expectations of all stakeholders—expectations that can only be met through unrealistic and ineffective amounts of multitasking. They reason that they are being unjustly tasked to fix a myriad of problems that seem to be rooted in situations over which they seem to have no control—especially the breakdown of the family and declining social values and morals. Moreover, these factors compound other problems teachers must address with students who arrive to their classrooms with disparate prior learning experiences and skills. Educators contend they do not have sufficient resources—especially time—to do more than teach an already overloaded academic curriculum upon which they are measured and for which they and their school systems are held accountable through standardized testing. Yet many stakeholders—especially those who hope to employ graduating students—also hold educators accountable for preparing students to communicate well, act responsibly, and work well with other people. How do standardized tests measure these attributes?

In other words, if the goal is to educate well, all students need to be prepared for life—to become productive and responsible citizens. In order to achieve that idealistic and worthy objective, educators must try to meet the needs of all stakeholders, especially the learning and behavior needs of *all* of their students.

On the other hand, educators must also be practical and realistic. Therefore, in order to educate well, they must work effectively within the limitations of existing resources. To do so requires educators to prioritize or set criteria for meeting needs with some likely being sacrificed. [Figure 26-1](#) presents a succinct definition of this core conflict that defines it *without* finger pointing.

Why is it so difficult to fix this problem in a way that does not compromise either existing resources or ensuring that all students become responsible and productive adults?

- Is it because we assume there is no way to teach life skills without sacrificing academic skills or vice versa?
- Is it because we assume actions to differentiate instruction to meet the learning needs of all students compromise resources beyond the breaking point?
- Is it because we assume students are unwilling or unable to take responsibility for their own learning and behaviors?

Is it possible to challenge and invalidate any of these assumptions? If so, what should be a solution and what should be the outcomes and other criteria to evaluate the solution's effectiveness?

What to Change to?

In the foundation of learning, a building block leads to a quality workforce and the future of a civilized society. That building block is the ability to think and communicate clearly. What if there were a set of concrete *thinking and communication tools* that could be used to teach prescribed curriculum in such a way that students:

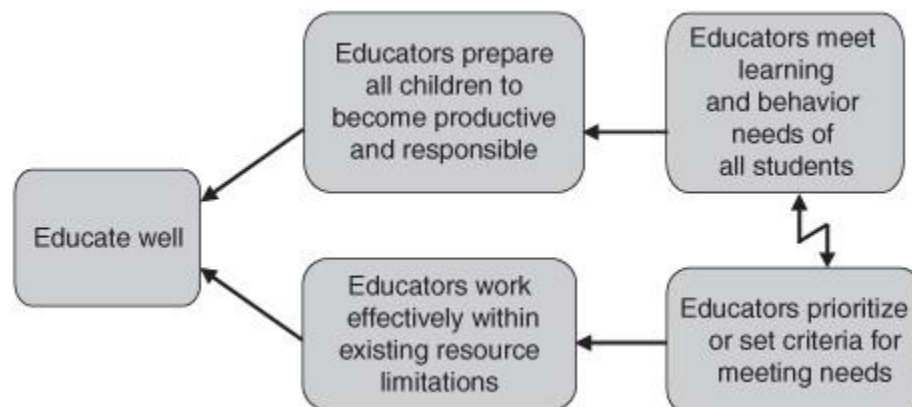


FIGURE 26-1 Core conflict. (Source: Kathy Suerken.)

- Develop their analytical thinking and communication skills at the same time,
- Apply the methods to problem solving and responsible decision making,
- Logically connect, interpret, and question information,
- Attain desired academic standards and benchmarks upon which they are measured,
- Perceive learning to be relevant, valuable, and transferable between subjects and real life, and
- Have the motivation and skills needed to feasibly achieve individual and collaborative goals?

Would these desirable effects not only prepare students to be productive and responsible but also enhance educators' existing resources, leaving them with more time for that which they consider most important and rewarding?

Of course, in order to achieve these outcomes and ensure they alleviate pressure on existing resources, the methodology of the tools must be simple, meet diverse student learning needs, and enable the learner to take ownership of solutions—whether they are in a textbook, playground, or a boardroom. If such tools and methodology to teach them actually existed, would educators use them? Let's consider the results of some of those who have.

How to Cause the Change?

TOCfE teaches three TOC thinking and communication tools that have graphic organizers and names: The Cloud, Logic Branch, and Ambitious Target Tree⁶ as depicted in [Fig. 26-2](#). These generic tools can be taught through applications

specific to curriculum delivery, behavior, and school management.

The Cloud

As we know, positive or negative effects in any one of these functions impact all the others. For example, when student behaviors improve, teachers are more able to focus limited resources on teaching and both of these outcomes help school leaders meet the needs and expectations of all school stakeholders. In other words, the whole system improves. A positive effect of successfully addressing the problem of bullying is one such example because the impact is felt not only by those explicitly involved in the bullying but also by all those indirectly affected.

Sometimes bullying manifests itself as name-calling. During recess in a Singapore elementary school, when Joel called Alex names and Alex reacted by using vulgar language and biting Joel on his arm, both nine-year-olds were sent to Vice Principal Wong Siew Shan's office. In a documented presentation (2000) to the 4th TOCfE International Conference⁷ Wong shared that her traditional response would have been to handle the problem for the children and then file a case sheet in the student's "misbehavior file" for future reference. A few days prior to the incident, however, she had taken TOCfE training sponsored by the National Institute of Education at Nanyang Technological University and was now looking forward to the opportunity of testing the TOC thinking tool, the Cloud, as a way of working through a problem by defining it through wants, needs, and a goal. [Figure 26-3](#) depicts the results.

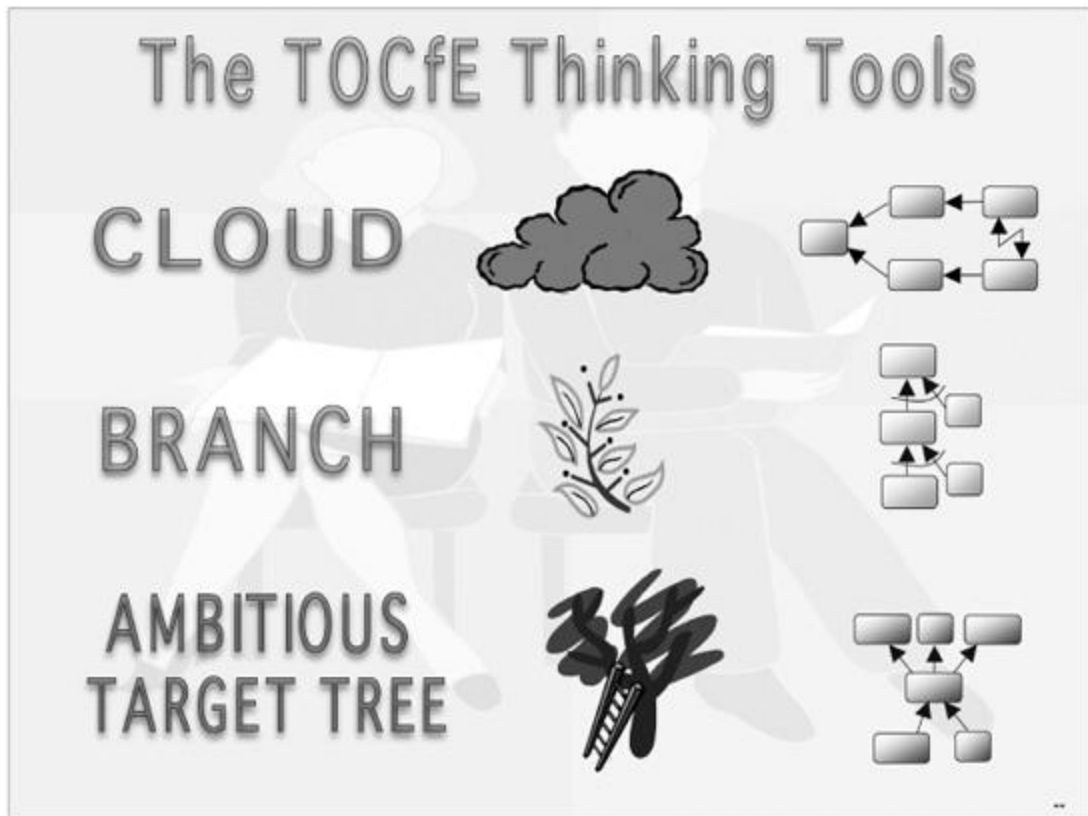


FIGURE 26-2 TOCfE thinking processes. (Graphics by Rami Goldratt. Source: TOCfE, used with permission.)

“It was heartening to note how easily they got the hang of how to use the Cloud template,” Wong noted. “After writing that his need was to have fun and, in order to do so, Joel wanted to call Alex names, Joel looked at me sheepishly and said that it wasn’t really true.”

The TOC tool guides students to see that, many times, their actions that lead to conflict are not based on clear thinking or accurate assumptions.⁸ The TOC process to explain the underlying reasons or assumptions why we take actions in order to get what we need is very effective to enable students to identify for themselves why sometimes their actions may not be appropriate and for *them* to create new and more responsible choices.

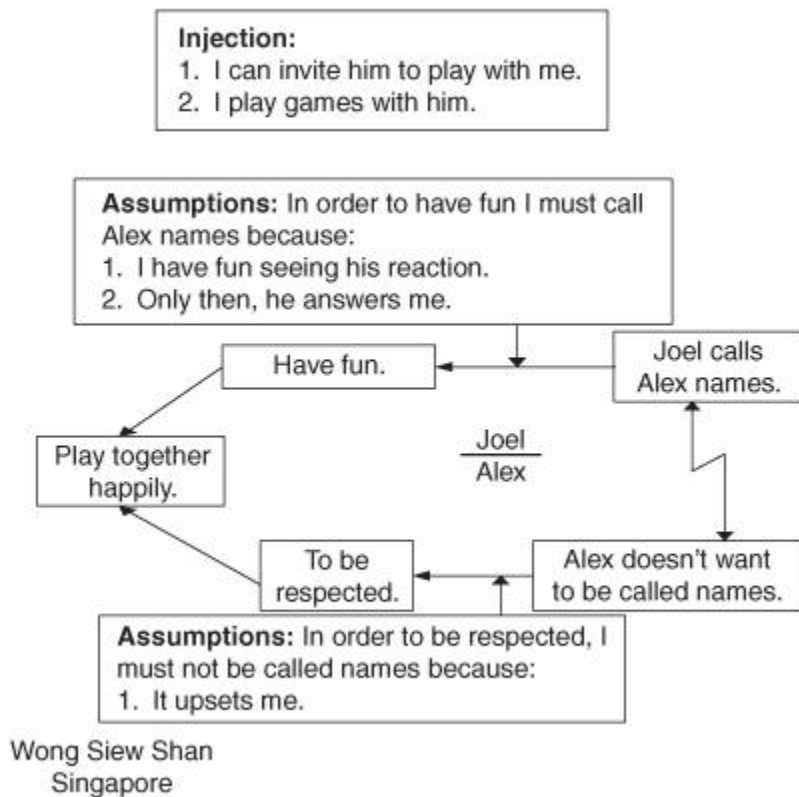


FIGURE 26-3 Name calling Cloud. (Source: TOCfE, used with permission.)

As Wong continues, “On surfacing his assumptions, Joel himself saw that they did not stand up to scrutiny. In fact, he came up with his own solutions and said that another way to meet his need to have fun would be to invite Alex to play with him.”

Additionally Wong pointed out that Joel also understood Alex’s need to be respected. Acknowledging and legitimizing the other side’s need in a conflict not only develops empathy but also a perspective well described by the words of (then) 13-year-old Niceville, Florida, student, Theresa Meyer: “The cloud makes you realize it is the situation that is the problem, not the people.”

The negative impact of name calling and bullying becomes exponential on a school campus when there are groups of students bullying other groups of students. As a

student assistant coordinator at a large Michigan high school, Doug Roby (1999) used the Cloud to resolve a situation involving seniors who were hazing freshmen or other new students [Fig.26-4]. In his words, “by hazing I mean they were trying to make them do ridiculous, humiliating, or even painful things. I used a Cloud in a group intervention . . . with about 20 senior girls on hazing. Within 30 minutes I explained the concept of the Cloud to the students, had them raise assumptions on one side of the cloud and come up with their own solutions. What a powerful tool to get students to really understand why they are doing something, what effect their actions have on others and to find alternative ways to meet their own needs.”⁹

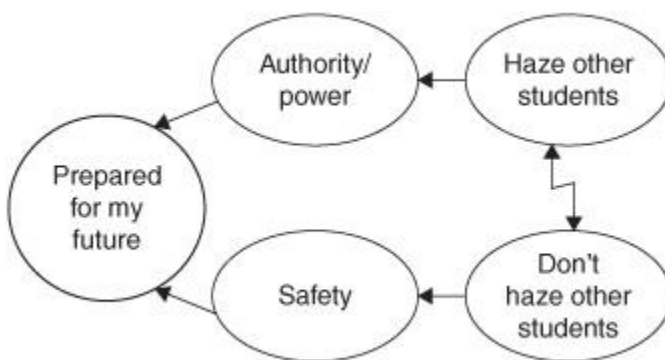


FIGURE 26-4 Group bullying Cloud. (Source: Doug Roby, used with permission.)

In describing a wider range of discipline issues at the school, Roby’s (then) Vice Principal, Ben Walker, noted, “Detentions, suspensions and, in one case, expulsion from the school only seemed to bring a temporary halt to the problem. After we started using TOC Peer Mediation, we were able to get to the root causes such as fear, jealousy, etc. As these students grew in self-awareness, they no longer felt a need to harass others. I find the drop in these cases remarkable.”¹⁰

This application of TOC to Peer Mediation has spread to schools in other countries—most notably to Colombia where, in 2005, then 15-year-old Ana Maria Conde and a group of her peers representing a TOCfE sponsored youth organization, AGOAL Academy, participated in a competition sponsored by the Universidad Nacional and the Mayor of Bogota. Ana and her team were required to submit a project that would present a well-defined problem, a concrete solution, and an implementation plan to achieve the solution. Of the 180 submitted projects, 36 were chosen to be presented in front of the Mayor and representatives of the University. As an award to AGOAL Academy for achieving first place on their use of TOC in Peer Mediation, the University sponsored TOC training of 10,000 students and 100 peer mediators.¹¹

The Cloud works with children of all ages to develop their abilities to solve problems wherever they encounter them. Therefore, in addition to painting a Cloud template on the playground in Nottingham, England for her students to resolve external conflicts during recess as pictured in [Fig. 26-5](#), then head teacher Linda Trapnell¹² in 1998 began to use Clouds to analyze problems in literature.

After reading an age-appropriate version of *Oliver Twist* to an assembly of 200 children between the ages of 4 and 7, Trapnell used the TOC processes to guide the students to define Oliver's internal conflict regarding peer pressure to steal. In TOC, a problem is not defined until it is presented as a conflict between two things. According to these young children, the conflicting choice was to be a pickpocket or not to be a pickpocket as noted in [Fig. 26-6](#). After summarizing the problem through the TOC graphic organizer, the Cloud, Trapnell then asked the students to think of reasons why, in order to satisfy his need for money, Oliver assumed he had to become a pickpocket.



FIGURE 26-5 Cloud on the playground. (Source: Linda Trapnell, used with permission.)

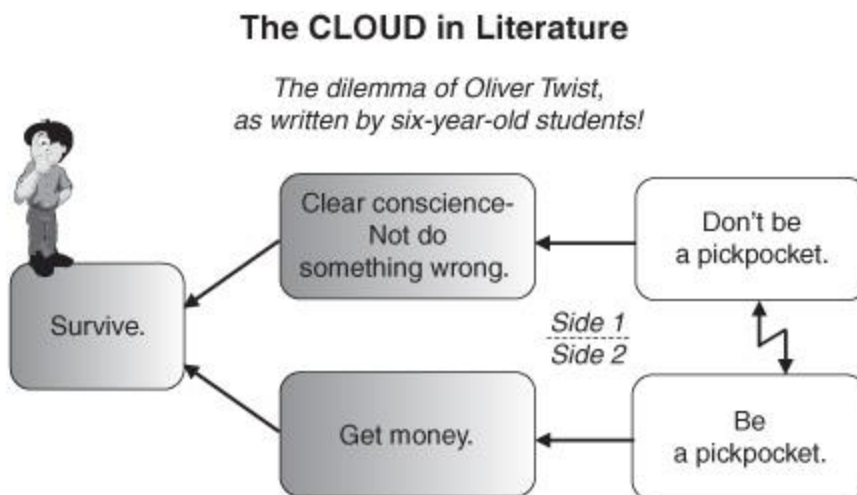


FIGURE 26-6 Cloud in literature example. (Source: TOCfE, used with permission.)

These reasons represent inferences¹³ and are an academic benchmark necessary to interpret information and to develop higher order thinking and problem-solving skills. Many strategies rely on combinations of definitions, examples, and visual illustrations to teach the concept of inference and how to apply it, but while helpful, they do not always sufficiently evoke the assumptions from which inferences can be drawn. The systematic, concrete questioning technique in the Cloud to raise assumptions is very simple and effective in enabling even very young students to draw inferences based on their individual experiences, knowledge, and opinions and to synthesize this information as they very simply explain the logical connections in information.

In this way, students are able to create their own scaffolds between their existing prior knowledge and the desired new knowledge. This scaffold also makes the learning more personally relevant to the students, thereby enhancing their motivation to learn. Enabling students to summarize, draw inferences, and identify deeper and broader perspectives of all sides are important academic benchmarks upon which students are tested. The more students are able to achieve these learning objectives for themselves through a systematic teaching methodology, the more they are able to meet their own learning needs.

After Trapnell's young students hypothesized that Oliver must have thought there was no way to acquire money other than by stealing, they became engaged in the next step of the process: creative problem solving. Guided by the TOC approach to find win-win solutions that meet both needs in the Cloud—in this case, the need for money and to maintain a good conscience—they created new solutions, such as Oliver could wash windows or get a job in a shop.¹⁴

Teacher-directed discussion on the assumptions and inferences that connect elements of the Cloud enables

students to be exposed to similar and different interpretations in a way that helps them evaluate and learn from their own and other perspectives. Therefore, this process also exposes gaps in understanding due to incorrect assumptions and inferences as when one student suggested Oliver could wash cars as a way of making money. If students are exposed to the appropriate missing information, then they can challenge their own inferences, as did this very young student, who revised his solution accordingly to “Oliver could look after horses.” When students realize they have the tools and skills to fix their own mistakes and to solve their own problems, they feel more justifiably self confident and motivated to do it again.

The Logic Branch

Students do innately try to make sense of the world around them. Therefore, they struggle when they try to learn facts and ideas that are disconnected and seemingly unrelated. The TOC *Logic Branch* helps students to create these logical connections using cause and effect to organize, sequence, and explain information in a way that makes sense and can be more easily remembered and analyzed. When analyzing text through Logic Branches, students are able to connect and scaffold information in a way that helps them derive and discover for themselves main ideas, generalizations, and other conclusions intended as lesson objectives. In this way, students are able to remember information more easily through the connections rather than having to memorize them as isolated facts. [Figure 26-7](#) illustrates how students are using the logic branch to connect information in a science lesson in Israel.¹⁵

In Tacoma, Maryland, 8th grade history teacher Manfred Smith (2007) found the Logic Branch highly effective to differentiate instruction to students of vastly disparate

levels of prior knowledge and skills. In a presentation at the 10th TOCfE International Conference,¹⁶ he reported that during yearly formal certification processes at his school, teams of evaluators could not distinguish between the work of his students considered to have learning disabilities and that of his students considered to be gifted. In the words of Jennifer Harris (2003), 8th grade Inclusion Teacher for World Studies, “. . . the TOC process has helped the students put an immense amount of facts and information into a logical and systematic order. From this, they are able to extract and apply information to writing prompts, group discussions, and expand their answers beyond basic recall. This is phenomenal because many of the students being served in this class were once self-contained special needs students who are reading at or near a third or fourth grade reading level.”¹⁷ The work of these students validates their capabilities to use a logical structure and methodology that enables them to make sense of—and explain—information at their own developmental level.

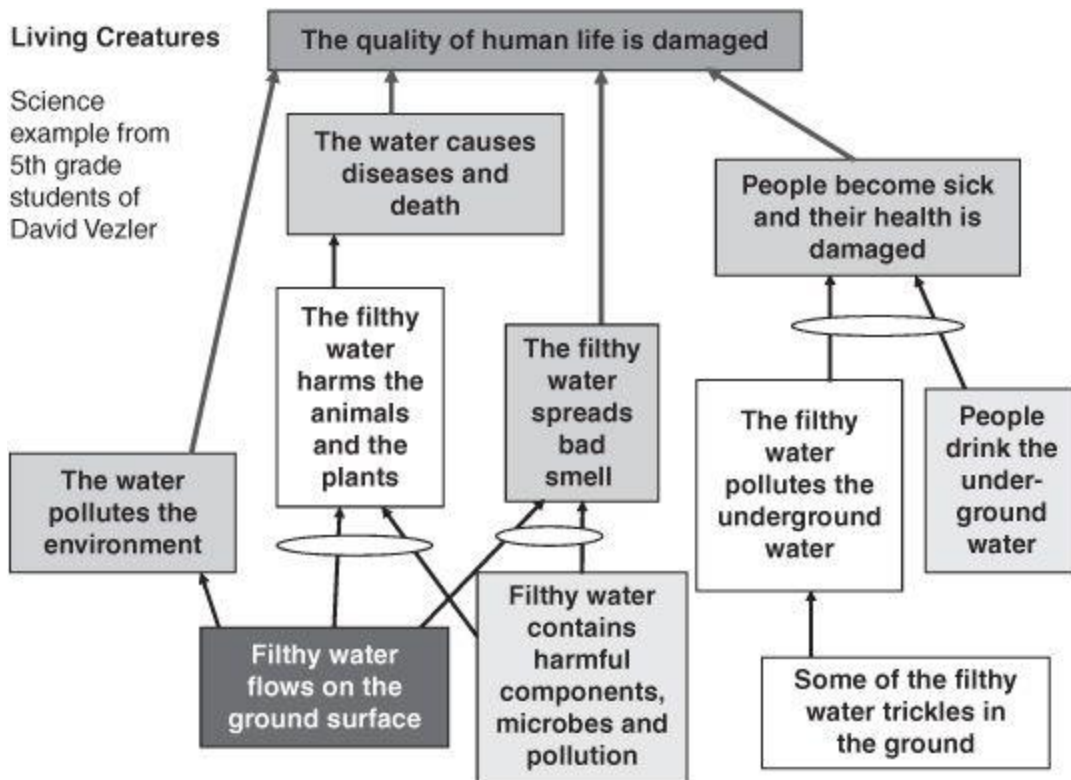
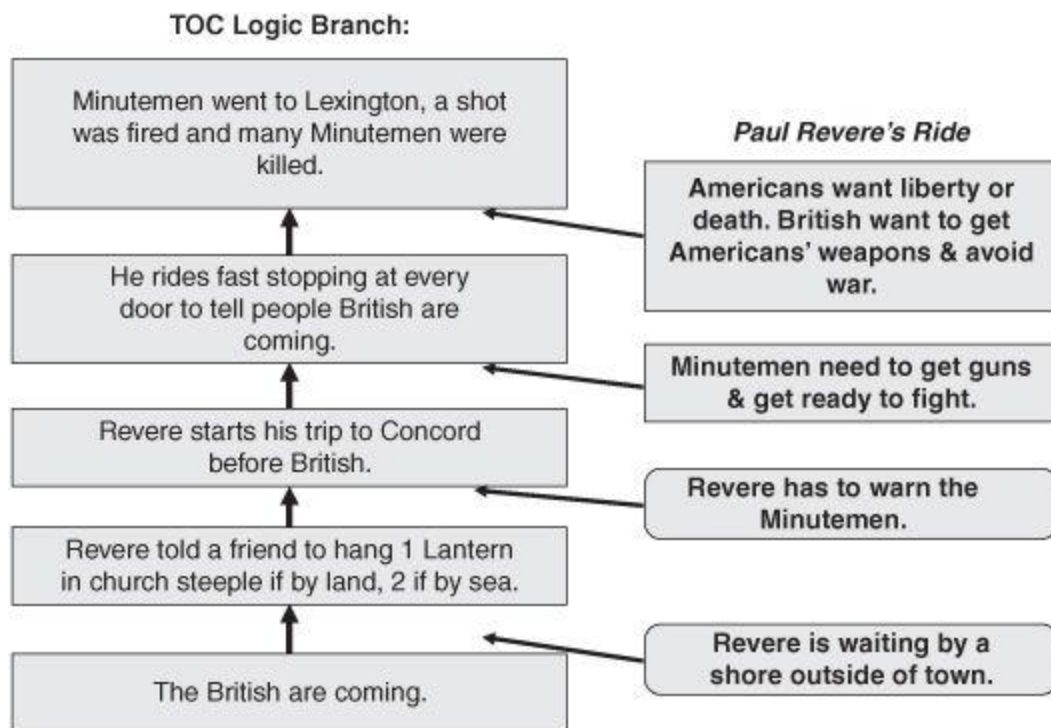


FIGURE 26-7 Logic Branch in science example. (Source: Gila Glatter, used with permission.)

An example of home educator Marilyn Garcia (2006) adds validity to this conclusion. She engaged her own 6- and 9-year-old children in the same history lesson because they both were able to contribute in a meaningful and focused way to lesson objectives by using the Logic Branch. After reading a poem to them about Paul Revere's ride, Garcia asked her younger child to write down the sequence of the main events by very simply prompting her with "what happened then?" Afterward, she asked the older child to provide supportive details and inferences that logically explained the chain of events also by using a simple questioning prompt of "if, then, because?" between the statements. The results presented in [Fig. 26-8](#) demonstrate that, using the same systematic thinking tool, these two students, within the same family and with very diverse

developmental skills and prior knowledge, were able to participate in a collaborative, focused, and developmentally appropriate way to achieve lesson objectives. [18](#)

The Branch, like the Cloud, can also be applied by children with diverse problems and in developmentally appropriate ways as a methodology to improve their relationships with others and by that improve their everyday lives. One of the first teachers taking a TOC seminar was Florida English teacher Belinda Small, who discovered that the Branch could very simply enable students to self-regulate their behaviors. She demonstrated that, when children can identify for themselves the cause-and-effect relationships between actions and consequences that affect them negatively, they are much more likely to take corrective actions on their own and even to establish different behavior patterns that lead to positive, rather than negative, outcomes.



Example of the children of Home Educator, Marilyn Garcia

FIGURE 26-8 Differentiating instruction with the Logic Branch. (Source: Marilyn Garcia, used with permission.)

Small writes, “Shortly after I was trained in TOC, I began to adapt one of the thinking methods (the negative branch) to get students to write down for themselves the consequences of their actions. The application was so effective with my 7th grade that soon all the teachers on my team began to send their disruptive students to me rather than the office because the process I was using is so effective! The amazing thing is that the students actually fix their own problems. All I do is get them to use the process. I think the students can write this so easily because they have experienced the chain of events. In this way, they are also developing a skill—cause and effect—which is sometimes otherwise difficult to teach. Using this method they can develop the skill by building—“scaffolding”—on prior knowledge rather than having to learn it as an independent skill.”

In describing the circumstances of a case study depicted in [Fig. 26-9](#), Small writes, “In one situation, when a student had been making disruptive noises in another teacher’s class, she asked for my help. The TOC Thinking Processes enabled this problematic student to *think for himself* the cause and effect outcomes of his actions. Although I did the initial writing of his words, at one point I had to leave to attend to my own class (obvious in the graphic). Nevertheless, this normally very disruptive student picked up the pencil—and the responsibility—and continued in his own words and graphics. We discussed what he could do to prevent the final outcomes and he wrote down some suggestions that were not new ideas. What was new in this case was that this time they were his ideas.”

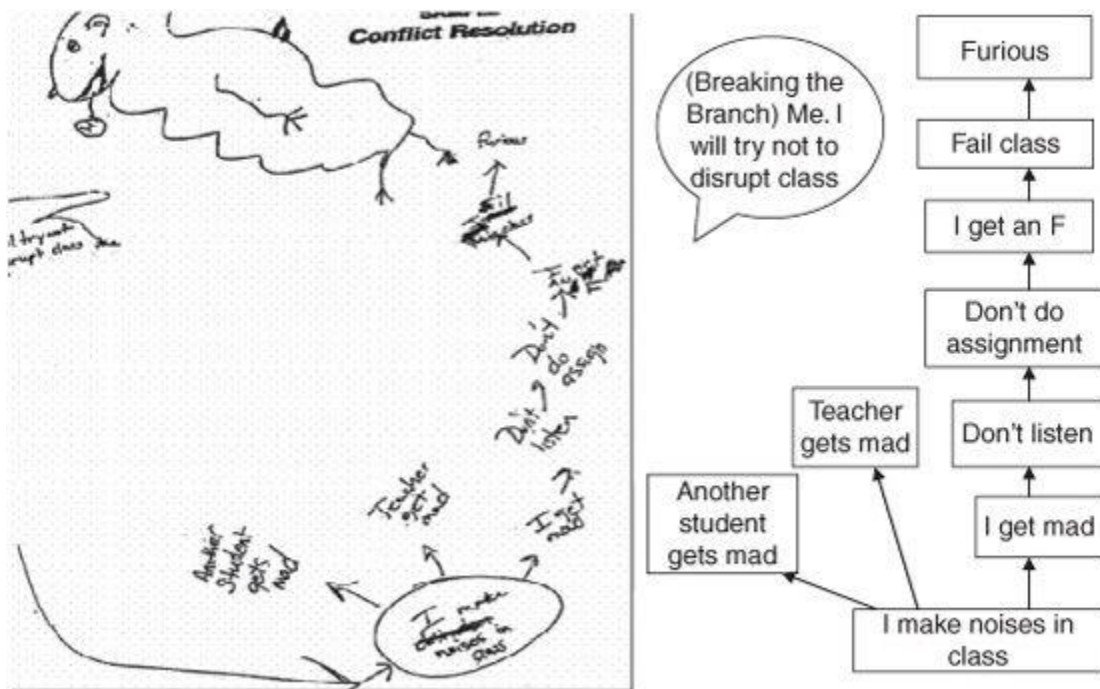


FIGURE 26-9 Using the Logic Branch with disruptive students. (Source: Belinda Small and TOCfE, used with permission.)

“The results? Although this student had been sent to the principal’s office 40 times in the previous 6 weeks, after this experience with TOC, he completed the rest of the school year (6 months) without a repeat offense with this teacher.”¹⁹

Holly Hoover of Virginia similarly quantified outcomes. “Of all my students who completed their negative [logic] branches on being tardy to class, none have been late again. 100% success! I like those odds! Not only do they see the consequences of their behavior from all angles (and where the behavior can lead to down the road) they also actually seem to like the assignment. Because of this, and the fact that it is not ‘writing sentences,’ the traditional assignment, the negative branches are always a ‘positive’ experience.”²⁰ Indeed, many teachers have students write the positive results of desirable actions, such as doing

homework, so that they can identify and take ownership of responsible choices that lead to positive outcomes for all concerned.²¹

Small's very simple, innovative, and highly effective application of the branch to children's behavior has been used with millions of children worldwide. In Perak, Malaysia, Principal Hajah Ahmad Rashidi uses a kinesthetic application of it by drawing templates for branches on the playground for children to use as hopscotch of cause-and-effect consequences as shown in [Fig. 26-10](#).²²

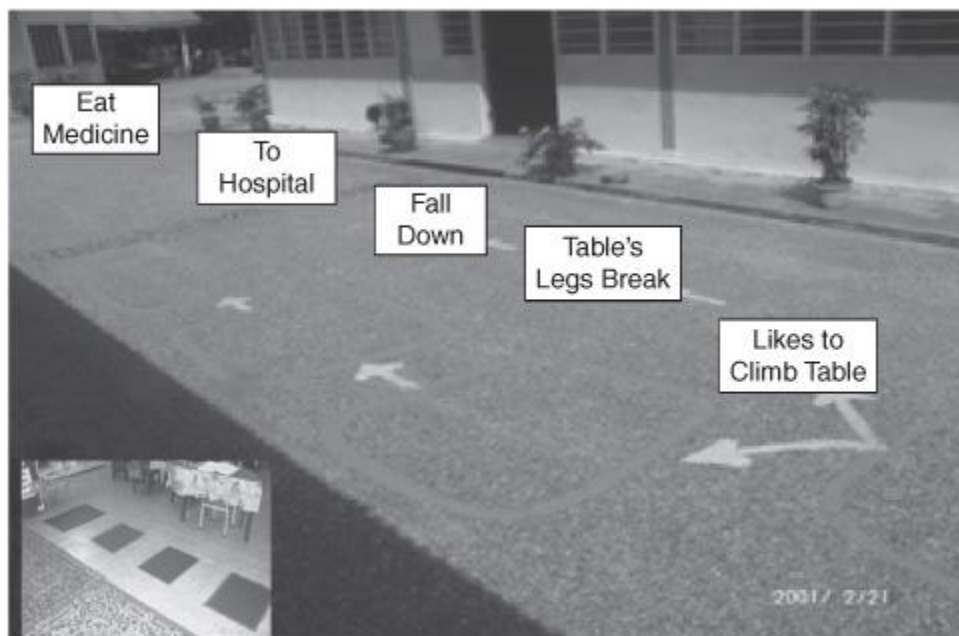


FIGURE 26-10 Using the Logic Branch as hopscotch. (Photograph and translation by Khaw Choon Ean, used with permission.)

TOCfE was introduced in Malaysia in 2000 through the Curriculum Development Centre as part of a Ministry of Education project called the Transition Program. The program was developed to address the problem of students who enter school at age 7 with different levels of readiness. In addition to the rich diversity of language within the

student population,²³ early childhood education before the age of 7 is at the discretion of parents and is not publicly funded.

Khaw Choon Ean, then head of Special Projects, designed the materials and engineered a cascade of training for all first grade/year one teachers in Malaysia—30,000 of them in 8000 primary schools and all within a mere 3 months. It was reported in Ministry tracking and review of the program that, even when introduced through curriculum lessons, students began to apply the TOC tools to real-life problems with siblings and classmates.²⁴ The use of the Cloud and Logic Branch has spread into Malaysian secondary education as a methodology to make instruction in social sciences more relevant and interesting. They have also been incorporated into civics textbooks at several grade levels as a methodology to promote responsible citizenship.²⁵

The Ambitious Target Tree

The results of a third TOC tool, the Ambitious Target Tree, further substantiates why the TOC Thinking Processes work with students to take responsibility for their own learning and behaviors. After first articulating a goal or “ambitious target,” the next step in the process requires students to analyze the situation before deciding on a course of actions—as do the other TOC tools. Therefore, students first identify “what to change,” which are the obstacles that prevent the attainment of the target. This is followed by “what to change to”—the intermediate steps that will remove the existence of the obstacle. “How to cause the change” requires that the intermediate steps are concrete and feasible actions that are properly sequenced. The process can be used to learn subject matter through the analysis of targets, obstacles, and intermediate objectives

or on individual or group targets such as one used at Maria E. Villarreal Primary School in Escobedo, Mexico.

<i>Be the best students</i>	
<i>Obstacles</i>	<i>Objectives</i>
Grumpy teachers	We listen to the teachers
Lazy students	We are prepared for class
We don't study	We study continually
Missed classes	We attend school regularly
We talk in class	We listen to the teachers
We bother classmates	We respect each other in class
We get to class late	We are on time for class
We do not participate	We gladly participate

FIGURE 26-11 Be the best students Ambitious Target. (Translation by Alexandrina Gonzalez. Source: TOCfE, used with permission.)

Teachers Zulema Almaguer and Miquel Perez Reyes used the Ambitious Target Tree tool as, in their words, “one of several TOC tools with very problematic groups of students to change their attitudes. In one case we worked with a group on the Ambitious Target of ‘Being the Best Students.’ When the students wrote their obstacles, they blamed others, but when they thought of ways to overcome their obstacles, they took the responsibility for the solution.”

As evident in [Fig. 26-11](#), in the first obstacle, the students characterized the teachers as “grumpy.” However, even though only of primary school age, these children were able to infer the reason that their behaviors might be

contributing to the teacher’s behavior and, from that inference, realize that they themselves could remove this obstacle through their own actions. The teachers conclude, “The students are learning to value themselves. The group was very much in conflict, but now I can see they are growing up because they are using the TOC tools to think through their problems.”²⁶

Does it make sense that most children are more motivated to implement a plan or project when they can meaningfully contribute to it? Using the tool in group projects not only engenders focused collaboration but also can expose obstacles that otherwise could be undetected and therefore continue to block the target. This was the situation with Florida teacher Belinda Small, working with TOCfE Senior Research Scholar Dr. Danilo Sirias from Saginaw Valley State University, Michigan.²⁷ Small applied the Ambitious Target Tree with her 7th grade English class on a subject highly relevant to those impacted by standardized tests.

Students of Belinda Small
(partial list)

Target: Raise Reading Test Scores		
Obstacles	Objective	Plan
1. The test is too long.	1. Make it shorter.	1. Use a pencil to divide passages into smaller parts.
2. I get stuck and can't remember the first paragraph.	2. Have reminders in the margins.	2. Summarize after each section/underline.
3. All the answers look the same.	3. Know the differences between choices.	3. Underline key differences in possible choices.

FIGURE 26-12 Raising reading test scores. (Source: Belinda Small, used with permission.)

Not surprisingly, when students suggested obstacles to a target of “Raising Reading Test Scores,” Small noted that many of them related to lack of confidence in standardized test taking, as noted in [Fig. 27-12](#). What was unexpected to her was why. When students verbalized an obstacle as, “All the answers look the same,” Small became aware that many students were primarily having trouble interpreting the multiple-choice answers, and that they lacked a strategy and specific actions to differentiate between the choices. According to Small, using the Ambitious Target tool enabled the students to develop their own strategy and tactics. In her words, “The tool enabled the students to create a step-by-step pattern to answer what to look for and do when reading questions and answers. This method enabled the STUDENTS to:

- think of the solutions
- create the language
- use THEIR logic
- form the connections between the State Academic Standards
- make the connections between the State Academic Standards and the FCAT²⁸ test questions.”

“Best of all,” she concludes, “they used it during the test. I felt the process had a big impact using very little time. It took about 30 minutes on one day to raise obstacles to the target. The next day we used about 15 minutes to think of intermediate objectives and another 30 minutes to organize the sequence of the objectives.”²⁹

The TOC thinking and communication tools provide a structure and the questions to empower students to analyze, derive relevance from, and apply what they are learning to their lives now and in the future. When children

have ownership not just of the answers but also of the questions that enable them to make sense of the world around them, they are much more able and motivated to take responsibility for what they learn and how they behave. This reality substantially fulfills stakeholders' expectations of good education in preparing children to become productive in the workplace and responsible citizens in a way that actually enhances the resources of those providing education—especially the resource of their time.

Yes, but . . .? How do we ensure that these results do not stagnate or deteriorate but endure and even progress? And what will be the impact of a progression of good results on our existing resources? Full circle . . . or a spiral?

A Process of Ongoing Improvement


There is nothing permanent but change—Herodotus
When students—or anyone—exhibit clear thinking, motivation, and improved performance, usually it is noticed, encouraged, and rewarded. While such success brings initial satisfaction and a justifiably enhanced self confidence, it can also create negative branches and raise new obstacles as conveyed in the words of Walt Whitman, “From every fruition of success, no matter what, shall come forth something to make a greater struggle necessary.” These obstacles can include:

- Rising expectations
- More work
- Ever changing realities

All of which can put pressure once more on our resources. Therefore, we need a *process* of ongoing improvement. In

TOC, the questions of change, just like the tools themselves, are not a one-time fix but, instead, systematically repeated as needed:

What to Change?
What to Change to?
How to Cause the Change?



The repeating cyclical applications of these questions and the TOC tools are intended to create spirals of ever-flourishing improvements whether in a person, a classroom, or an organization—all of which combine in TOCfE. Therefore, not surprisingly, TOCfE has experienced the same core conflict as in [Fig. 26-1](#) and the need to revisit the three questions, given the phenomenon of rising expectations and changing needs from existing and new, more diverse TOCfE stakeholders.

As noted, one of the strengths of the TOC tools is that they can be taught and made relevant in classrooms—and with other groups—of people who have very divergent levels of knowledge, skills, and interests. This relevance has led to diversification within the TOCfE network in terms of specialized applications and interventions, particularly to enable those with special behavioral needs. For example, in The Netherlands, a TOCfE consultant, Fiet Muris, is using TOC with groups of children and parents who are part of the Romani population. They live in caravans and feel the sting of isolation, prejudice, and low academic achievement of these children who attend local schools. The TOC intervention is so effective that local schools and supportive government agencies have joined in creating solutions that are beneficial to all concerned.³⁰

Other such specialized TOCfE applications include:

- Children and adults who have dyslexia
- Children considered to be gifted

- Children at risk to develop addictive behaviors
- Students in dropout prevention programs
- Children diagnosed with Down syndrome and cerebral palsy
- Children with Asperger syndrome
- Children considered to be asocial and with other significant behavior disorders
- Adult inmates in the penal system³¹

The TOC interventions with all these special interest groups have been very effective, as evidenced by the many case studies presented in TOCfE conferences and posted to the TOCfE Website, as well as by a growing amount of research. The research of Edyta Sinacka-Kubik (2006-2007), a PhD student at the Psychology Institute at the University of Gdansk, Poland, who was trained in the tools in a 3-day seminar in 2006, is one such example. Her stated hypothesis is: “There is a possibility to overcome school-educational difficulties when it comes to asocial children by applying the TOC for Education support program.” The research involved an experimental group consisting of 22 children regularly attending four sociotherapeutic centers; the implementation of the 18-month TOCfE project; and regular meetings at least once a week for 1.5 hours. The control group contained 22 children regularly attending four sociotherapeutic centers.

Some of her findings are presented in [Fig. 26-13](#) and [Fig. 26-14](#).³² These results are considered statistically significant and they include this summary presented at the 10th TOCfE Conference in Fort Walton Beach, Florida:

- “TOC group gained significantly lower results in *Antisocial Behavior Scale* after the experiment.
- TOC group gained significantly lower results in *Withdrawal Scale* after the experiment.

- TOC group gained significantly higher results in *Socialization Scale* after the experiment.
- TOC group gained significantly higher results in *Motivation to Learning Scale* after the experiment.
- TOC group made much *bigger progress* than control group did during the experiment.” (PowerPoint summary)

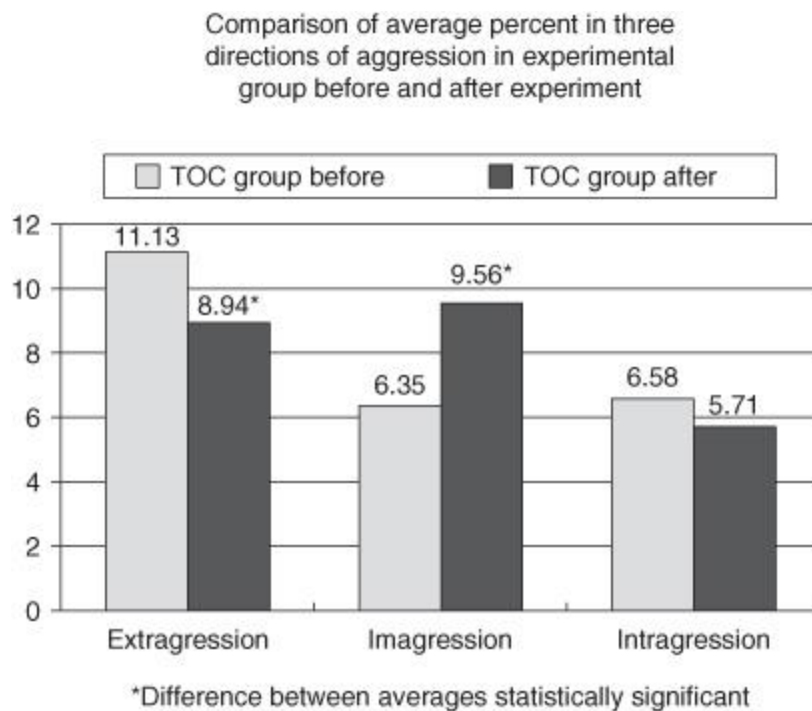


FIGURE 26-13 Comparison of aggression research. (Graphics by Edyta Sinacka -Kubik. Source: TOCfE, used with permission.)

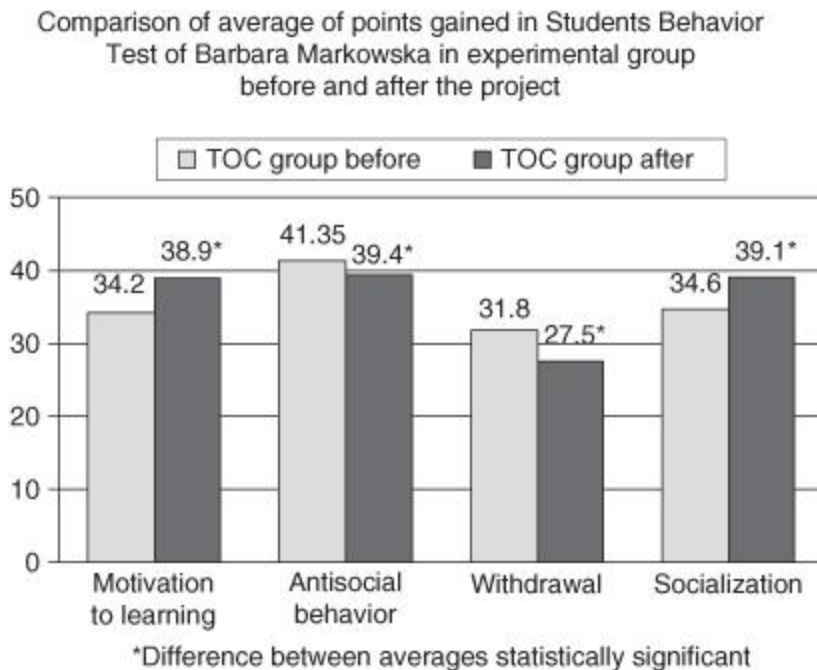


FIGURE 26-14 Antisocial behavior research. (Graphics by Edytak Sinacka-Kubik. Source: TOCfE, used with permission.)

Sinacka-Kubik concludes, “On the grounds of these optimistic results, visible even in a small group, we can think that in much bigger groups, the effect would be more significant. This research has encouraged us to start a new, much wider project.”³³

Improvements in not only communication and behavioral skills but also in performance were validated in the PhD research of Dr. Jenilyn Corpuz, principal of a high school of more than 3600 students in Quezon City, Philippines. Research for her dissertation looked at “The Impact of the TOC Tools to Determine the Effects of the Theory of Constraints for Education (TOCfE) Tools as Intervention Instruments in the Teaching-Learning Processes in Technology and Livelihood Education.” The study, as shared at the 8th TOCfE International Conference,³⁴ included four second-year homogeneous classes at New Era High School that have an average of 60 students per class, handled by one teacher.

One of the specific objectives of the research was to assess the performances of students in experimental and control groups in terms of self-efficacy in communication and behavioral skills. Another specific objective was to evaluate if there was a significant difference between the performances of students in experiment and control groups of content pre-assessment and content post-assessment. Student outputs were rated using a rubric based on the basic education curriculum and TOCfE concepts. Notable increases of the percentage mean scores in experimental group are indicated in the results.³⁵

This published research is helping to address TOCfE's initial lack of empirical evidence and the justifiable need to demonstrate that the TOC tools are aligned with sound learning theory and research methods. More international research projects are now underway in Israel, the United Kingdom, and the United States to test implications of the use of TOC to enhance emotional intelligence and the delivery of science and mathematics curricula. This ongoing research reflects the progression of TOC's continuous improvement by using the three questions to know where to focus improvement efforts.

A particular area of focus has always been materials. A comprehensive series of self-learning workbooks for children of various ages was created in the late 1990s and early 2000s in Israel under the direction of Gila Glatter while teaching at Talpiot Teachers' College, Tel Aviv, Israel.³⁶ Some of these workbooks are available in English as is a story that teaches the Cloud for middle school children. A CD-ROM is also available³⁷ that can be used to teach all three tools through an animated children's story. Entitled "The Story of Yani's Goal," this piece of literature has a moral: You can achieve your goals in life if you think your problems through to win-win solutions. The story, soon to be released in book format, can be used in reading classes along with a

teacher's guide designed to enhance reading comprehension skills through the content of the story.

TOCfE began training in 1995 with TOC materials³⁸ written for business and industry—particularly those developed for human behavior and based on Eli Goldratt's (1994) book, *It's Not Luck*. In order to make these workbooks user friendly and more relevant for educators, the workbooks were carefully adapted³⁹ and the TOCfE training materials for seminars are entitled "TACT" (Thinking And Communication Tools). Available in Spanish, Dutch, Hebrew, Russian, Serbian, Portuguese, Polish, and English, these materials primarily teach the three tools through behavior applications. Therefore, much of the dissemination and diversification has been to apply the processes in counseling and interventions with children—and adults—with special behavioral needs.

As noted, Corpuz's (2005) research was based on teaching the tools as interventions to improve cognitive performance as was the Master's thesis of Adora Teano (2005-2006), who concluded, "TOCfE thinking tools have significant effect on the improvement of the grades of the students in English I."⁴⁰ Action research case studies from the trenches also substantiate the "proof in the curriculum pudding" that the TOC tools, once learned, enable students to create the scaffolds needed to meet their own learning needs, thus saving time and other resources both in and out of school. They are interventions, however, rather than preventive strategies and require the teacher to transfer the application from behavior to that of curriculum. Because of the need to work within existing resources, teachers need the transfer of application to be from existing curriculum to behavior so that students learn these needed life skills while being taught existing curriculum. This latter approach is very much aligned with conclusions for implementing change in educational systems drawn by Dr. Audrey Taylor

(2002, 126) in her PhD research that examined TOCfE concerning change agents and performance measurements. These include:

- “Successful change is accelerated when the training is application specific so that the user can easily implement the new methodology.
- Regardless of the content of the change methodology, the faster the results in the classroom, the faster the dispersion of the new methodology.”

Therefore, as TOCfE strives to improve, a “what to change” has been to refocus resources on a new generation of workbooks and seminars. Written in 2009 and entitled, *Thinking Across the Curriculum*, they teach the generic tools specifically through curriculum applications and are written to meet standard professional development criteria with detailed, measurable learning objectives. These materials are being translated into Polish in support of a comprehensive training program beginning in December 2009 and sponsored by MSCDN, a Polish professional teacher training center,⁴¹ and the Polish National Institute of Psychological Support.

The focus on curriculum is also the catalyst for an international action-based research project that involves exchanging TOCfE-based curriculum lessons and student work through collaboration on the Internet. Initially the project will include schools in Israel, Mexico, and the Philippines where teachers are using the TOC tools in science, language arts, and mathematics. The project is being organized by schools that are using the TOC tools in counseling and management as well as in curriculum. This holistic approach to school improvement models the ideal to teach by example in every function of the school.

TOC tools, simple enough to be used by kindergartners



Alderman Pounder Infant and Nursery School United Kingdom

FIGURE 26-15 Ambitious Target of a wedding plan. (Source: Linda Trapnell, used with permission.)

Another identified tactic to bring strategic, ongoing improvement in TOCfE is to focus current fundraising efforts in support of creating and maintaining a cutting edge e-library that will house the global bank of examples, research, and other work that TOCfE practitioners wish to share. While protecting intellectual properties of those sharing work, it will provide opportunities for readers to learn from—and improve—ideas in a win-win way that tracks their contributions as well. Through these means, TOCfE can foster innovation and collaboration of all education stakeholders who want to make a meaningful difference by allowing others to use their work in a way that not only, as Jesse Hansen envisioned, “pays in kind” but also “pays it forward.”

The TOC tools work without regard to age, culture, or political system as reflected in the wealth of work and vision of all those who have so graciously shared it with TOCfE.

The extraordinary scope of TOCfE applications are well characterized in the words of former LASD⁴² School- and District-Based Administrator Denise Meyer: “The TOC tools are simple enough to be used by kindergartners and profound enough to be used by CEOs.”⁴³ [Figure 26-15](#) and [Fig. 26-16](#) illustrate this claim through an Ambitious Target of a United Kingdom nursery school student⁴⁴ and the Ambitious Target of (then) National Capital Regional Director of Philippine Department of Education, Sports and Culture (DECS), Dr. Cora Santiago, who supervised 17 school superintendents responsible for educating 8 million children.⁴⁵ Using the strategies identified through the Ambitious Target tool, Dr. Santiago wrote a TOC tactical logic branch entitled “ZERO NON-READER” and submitted it to the Asian Institute of Management (AIM) which accepted her as a scholar. After graduating, Dr. Santiago conferred with McDonalds which was interested in the proposal as a corporate responsibility project. The project, called “BRIGHT MINDS READ,” was initially implemented in the National Capital Region and is now one of the flagship programs in the country.

Ambitious Target of DECS-Manila for Year 2004 80% of Grades I-VI Pupils are independent readers in English in 2004	
Obstacles	Intermediate Objectives
1. Lack of well-trained supervisors, administrators and teachers in reading.	1. Supervisors, administrators, and teachers adequately trained.
2. No systematic and organized monitoring scheme.	2. Division-wide monitoring scheme installed, with MTs assisting in monitoring.
3. Funds not available.	3. Government funds released for initial program support.
4. Evaluation measures for progress reporting inadequate.	4. Mid-year and year-end oral and written tests administered in reading I-VI.
5.1 Insufficient quality books.	5.1 Quality books available: READ-A-THON and Battle of the Books are utilized.
5.2 Assessment material inadequate.	5.2 Authentic forms of assessment portfolio, journals, scales, checklists etc. utilized.
5.3 Dearth of instructional materials.	5.3 Varied instructional materials prepared.

6. Application of theories in IRPE in actual classes difficult.	6. Division, district, clusters and school level demonstration classes in I-VI held.
7. Model lesson plans inadequate.	7. Lesson plans available and production continued by MTs.
8. Lack of collaboration among Filipino and English supervisors and teachers, I-VI.	8. Brainstorming and regular evaluation sessions among Filipino and English supervisors and teachers, I-VI.
9. Teachers' oral and written communication skills in English inadequate.	9. School-based LAC sessions held to improve teachers' communication skills.
10. Poor communication skills (oral and written) among pupils I-VI.	10. Classroom interventions: flexible grouping, differentiated instructions, assessment and assignment given focusing on poor readers.
11. Parental support inadequate.	11. Parental support and involvement attained.
12. No division-wide integrated reading program in English to make each child a reader.	12. A Division Integrated Reading Program in English (IRPE) prepared and clearly disseminated.
13. Lack of coordinated support structures for sustenance to pursue programs till 2004.	13. Strategic plan for program sustenance till 2004, implemented, monitored, evaluated and modified in coordination and collaboration with all concerned.

FIGURE 26-16 Ambitious Target of Philippines DECS. (Source: TOCfE, used with permission.)

The reason for this breadth and depth of application is a common denominator that has roots deep enough to encompass all education stakeholders. The Socratic nature of these systematic, logical tools enables people to discover—for themselves—answers that make sense.

When these solutions are needed because of conflicts or to address negative consequences, children of all ages are able to take ownership of responsible choices without losing face. This “accountability with dignity” is sustainable because it is intrinsically created rather than extrinsically imposed.

The same Socratic tools work in a classroom to give children ownership of what they are learning. When students derive their own answers, it engages perhaps the most important ingredient in education: the student’s wish to learn.

As these simple, robust tools continue to combine with the collaboration of all who want to meaningfully touch the

future, TOCfE will enable more and more children around the world not only to become responsible and productive adults, but also to engage in everflourishing life-long learning that, like success, is not a destination but a journey.

References

- Almaguer, Z. M. and Reyes, M. A. 2001. "Changing the mindsets of groups of disruptive students." Presentation at 2001 TOC TOCfE Mexico Conference, Monterrey, Mexico (March). <http://www.tocforeducation.com/att-b/attb02.html>
- Anaya, J. De Ninos and de la Luz Pamanes, M. 2001. "Violence in the home." Presentation at 2001 TOCfE Mexico Conference, Monterrey, Mexico (March). <http://www.tocforeducation.com/cloud-b/cb23.html>
- Anonymous. 1994. "Applications of TOC by Okaloosa County Educators," (June 6-9) Jonah Summer Conference, Ft. Walton Beach, FL: Avraham Y. Goldratt Video Report Series.
- Conde, A. M. 2005. "AGOAL Academy." Presentation at 8th TOCFE Education International Conference, Seattle, WA. (August).
- Corpuz, J. 2005. Impact of the TOC tools to determine the effects of the Theory of Constraints for Education (TOCfE) tools as intervention instruments in the teaching-learning processes in technology and livelihood education. PhD dissertation, University of the Philippines.

- Corpuz, J. 2008. "Curriculum-based research projects." Presentation at 11th TOCFE International Conference, Warsaw, Poland (October).
- Garcia, M. 2006. "Differentiated Instruction." Presentation to Maryland Home Education Association, Columbia, MD (November).
- de Gaza Gonzalez, A. and Rodriguez, M. 2001. "Enabling juvenile offenders to set goals." Presentation at 2001 TOC TOCFE Mexico Conference, Monterrey, Mexico (March), <http://www.tocforeducation.com/attb/attb09.html>
- Glatter, G. and Kovalsky, S, 2000. *The Way of Achieving a Target*. Two workbooks (in Hebrew). Tel Aviv, Israel: TOC for Education Israel.
- Glatter, G., Wiess, N., and Talek, M. 1999. *Solving Day-To Day Conflicts*. Three workbooks (in Hebrew). Tel Aviv, Israel: TOC for Education Israel.
- Goldratt, E. M. 1984. *The Goal: Excellence in Manufacturing*. Croton-on-Hudson, NY: North River Press.
- Goldratt, E. M. 1994. *It's Not Luck*. Great Barrington, MA: North River Press.
- Grienbert, M., Goldratt, R., and Glatter, G. 2002. *The Rainbow in the Cloud* (in Hebrew). Tel Aviv, Israel: TOC for Education Israel.
- Harris, J. 2003. <http://www.tocforeducation.com/references/html>

- Hoover, H. 1999. "I have had no further problem with tardiness." <http://www.tocforeducation.com/branch-b/bb02.html>
- Hutchinson, M. "TOC in counseling: Taking responsibility for learning, a classroom behavior intervention." <http://www.tocforeducation.com/att-b/math3.html> and <http://www.tocforeducation.com/att-b/math4.html>
- Khaw, C. E. 2004. "TOCfE in Malaysia." Presentation at 7th TOCfE International Conference, Ft. Walton Beach, FL (May).
- Khaw, C. E. 2005. *Thinking Smart: You Are How You Think*. Selangor, Malaysia: Pelanduk Publications.
- Khaw, C. E. 2006. "100 children × 100 days × 100 clouds." Presentation at the 9th TOCfE International Conference, Leon, Mexico (Sept).
- Meyer, D. 1999. "TOC and the Children of Los Angeles." Presentation to 3rd International TOCfE Conference, Los Angeles, CA (Aug).
- Meyer, D. and Kelly-Weekes, R. 2000. *LA CREME: Los Angeles Conflict Resolution Education Model for Educators*. Los Angeles, CA: Los Angeles Unified School District.
- Muris, F. 2008. "TOC and the children of Romani populations." Presentation at the 11th TOCfE International Conference, Warsaw, Poland (October).
- Roby, D. 1999. "An alternative to hazing." <http://www.tocforeducation.com/cloud-b/cb2.html>

Santiago, C. and Visaya, L. 2002. "TOC and Literacy in the Philippines." Presentation at the 6th TOCfE International Conference, Nottingham, England (July).

Sinacka-Kubik, E. 2007. "How to take advantage of Theory of Constraints for Education program to support children's psychosocial development." Presentation at 10th Theory of Constraints for Education International Conference, Ft Walton Beach, FL (October).
<http://www.tocforeducation.com/researchlist.html>

Sirias, D., de Garza Gonzalez, R., Rodriguez, M., and Salazar, E. 2007. *Success: An Adventure*. Saginaw, MI: Author.

Small, B. 2003. "The case of the disruptive student." Presentation at the 7th Theory of Constraints for Education International Conference. Ft Walton Beach, FL (May), <http://www.tocforeducation.com/branch-b/bb01.html>

Smith, M. 2007. "Academic applications: Generating interest, knowledge, motivation and success through TOC thinking tools." Presentation at the 10th Theory of Constraints for Education International Conference, Ft. Walton Beach, FL (October).

Suerken, K. 2008. "The story of Yani's goal" (CD). Niceville, FL: TOC for Education, Inc.

Suerken, K. 2009. *Thinking across the Curriculum: The Cloud, the Logic Branch and the Ambitious Target Tree*. Three workbooks for teachers. Niceville, FL: TOC for Education, Inc.

- Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. The Theory of Constraints International Certification Organization Dictionary. <http://www.tocico.org/?page=dictionary>
- Taylor, A. G. 2002. An empirical investigation of the change agents and performance measurements effective in the diffusion of the Theory of Constraints for Education (TOCfE) and implications for business entities. PhD dissertation, Wayne State University.
- Teano, A. 2006. The effectiveness of integrating the Theory of Constraints for Education in the teaching learning processing for English I. Masters thesis. University of the Philippines.
- Trapnell, L. 1998. "From under a cloud," *Child Education* 75(August):46-7.
- Trapnell, L. 1999. "Storytelling: Oliver Twist." <http://www.tocforeducation.com/cloud-c/ccOLhtml>.
- Trapnell, L. 2000. "Learning how to make logical plans." <http://www.tocforeducation.com/att-c/attc04.html>
- Trapnell, L. 2003. "Case study two," *Primary Leadership Paper* 1:8,14-18.
- Trapnell, L. 2004. "Theory of Constraints: Thinking for a change," *Teaching Expertise* 2(Winter):35-37.
- Walker, B. 1997. Speech at North Central Association Commission on Accreditation and School Improvement Conference, Chicago, IL (April).
- Wiess, N. and Talek, M. 2001. *Think Before You Act*. Two workbooks (in Hebrew). Tel Aviv, Israel: TOC for

Education Israel.

Wong. S. 2000. "TOC mediation to stop name calling."
Presentation at the 4th Theory of Constraints for
Education International Conference, Monterrey,
Mexico (August),
<http://www.tocforeducation.com/cloud-b/cb7.html>

About the Author

Kathy Suerken has been President of TOC for Education, Inc. since it was established by Dr. Eliyahu M. Goldratt in 1995. During Kathy's leadership, TOC tools and concepts have been taught in 23 countries to well over 200,000 adults involved in education with an impact on more than 8 million children worldwide.

In addition to speeches at national and international education conferences—such as the National Educator's Congress of the Philippines and the 8th International Conference on Thinking—Kathy's business presentations include a keynote address at the APICS SIG Symposium. Kathy is author of "The Story of Yani's Goal," an animated novelette, numerous TOCfE seminar workbooks, and is co-author of " . . . the Never Ending Story," a children's workbook on conflict resolution.

A former middle school teacher with a BA in History from Wittenberg University, Kathy has received extensive business management training in TOC. She is an AGI Jonah and Jonah's Jonah (facilitator) and is certified in the TOC Thinking Processes by the TOC International Certification Organization. Kathy has received extensive experience in the art of teaching and learning from her students who range from children to Ministers of Education.

Kathy can be reached at Suerken@cox.net.

CHAPTER 27

Theory of Constraints in Prisons

Christina Cheng

Introduction

Theory of Constraints (TOC) has a distinguished history of success dealing with constraints in business and education. In Singapore¹ where math and science rankings are among the top in the world, its educated human capital is regarded as its most prized asset. At the same time, given the country's limited population, it is viewed as a core constraint. To exploit this constraint and in line with ongoing government policy to improve workforce productivity, an opportunity arose to help long-term unemployed workers reintegrate into the Singapore workforce using the TOC *Thinking Processes* (TP).

In August 2006, the *National Trade Union Congress* (NTUC) through its Job Re-creation Program in conjunction with the Rehabilitative Division of Singapore Prison Services engaged TOC Asia Pte Ltd. to help prepare pre-release adult prison inmates for outside employment using the TOC TP. As part of the pilot study, selected inmates would attend a TOC mindset management workshop immediately followed by a NTUC job fair at the end of October 2006 to help them secure a job before release. The end goal presented to TOC was to reduce the high job attrition rate of ex-inmates upon release. This meant that any behavioral or mindset change observed during the TOC workshop must be sustained

outside the relatively stable prison environment in the face of uncertain external influences in order for the project to be deemed successful.

It was clear that a formidable task was ahead. A major obstacle in the preparation of the workshop was a lack of student conformity in the pilot group, with regard to age, language, education, race, and type of offense. This resulted in extreme variations in class profiles. In a particular training session, a Malay-speaking elderly inmate, slightly deaf and illiterate, could be seen sitting next to an English-educated postgraduate sociologist! Coupled with no available generic training materials, limited prison intuition, no formal background in teaching or psychology to address the disparate range of chronic negative behavior, the biggest question was whether we could adequately address all individual training needs of the diverse pilot group.

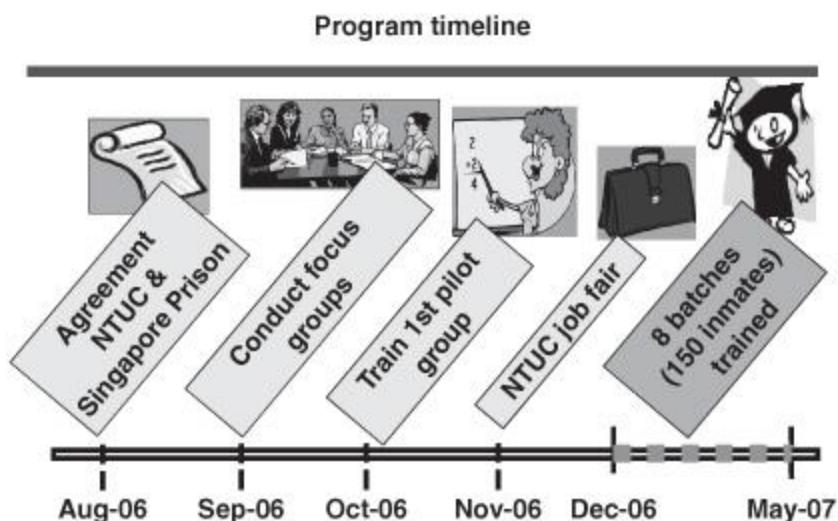


FIGURE 27-1 The project timeline.

The other challenging factor was the short course timeframe. In order to meet the October NTUC job fair schedule, the course duration was limited to 18 hours spread over six sessions over 2 weeks. Was it possible to change a person's mindset within such a short time? Project

success would be measured by the percentage increase in job retention over the first 3 months of employment upon release. Did it work? At the end of the pilot study,² job retention over the said period rose threefold from a historical 20 percent to an astonishing 59 percent. Numbers aside, however, we leave it to the reader to gauge the success of the overall project. The program timeline is provided in [Fig. 27-1](#).

What to Change?

Preliminary Study

Demographics of the pilot study group were as follows:

- 60 male adult offenders
- Age range 21 to 60 years
- Primary school education or below (46 percent), secondary (32 percent), pre-university GCE, N, and O level (12 percent), technical (9 percent), degree (1 percent)
- Malay (50 percent), Chinese (42 percent), Indian (7 percent), Other (1 percent)
- Weak to basic English comprehension

Independent focus group sessions were conducted with the prison officers and inmates prior to workshop commencement to better understand how TOC could be used to bridge the gap between existing rehabilitation and job preparation programs being conducted. Put more simply, we needed to know why the historical job attrition rate was so high. To promote open discussion, uniformed staff and inmates were interviewed separately using a simplified TOC

Prerequisite Tree (PRT) framework with an ambitious target of “To be successful in the workplace.”³

From the generic list of obstacles raised (e.g., my family does not accept or support me, people look down on me, I am easily influenced by my negative peers, I don’t know how or I am not ready to change, I have no positive role model), it became evident that while many inmates had acquired valuable technical and soft job skills (e.g., IT, communication, and interview skills) and underwent targeted rehabilitation to overcome specific types of negative behavior (e.g., anger management, drug addiction) during their prison stay, not all were mentally prepared to face society regardless of whether they had a ready job in hand.

For most, what was missing was the confidence and conviction that they could properly reintegrate into their family and the workforce upon release. This was not surprising, given that many had repeatedly tried and failed, leading to a history of multiple offenses. For the majority, prison was not an uncharted path. From an early age, they had unsuccessfully navigated through the maze of correctional facilities leading to their present situation. Many had come from dysfunctional home environments with little or no family support, leading to a heavy reliance on negative peers to provide a sense of identity and belonging. The resulting lack of positive role models provided a distorted sense of values and justification of what would normally be regarded as negative behavior.

For others, prison provided a false sense of security, away from the stress and pressures of daily life. In the wry words of one TOC participant, “All our major needs such as housing, food and medical, even new glasses, are provided for.” Many doubted their ability to survive financially without support from “easy money” gained from illicit activities and worried about their lack of willpower to withstand the

myriad of external social pressures needed for sustainable change. Despite a desire to change, they knew their personal limitations and battled inner demons to overcome familiar temptations as their release date drew near.

Even for those determined to start afresh, they questioned their likelihood of finding “good” friends to lead a normal life. Despite the prospect of a secure job to meet their basic monetary needs, this was overshadowed by internal fears that they would be stigmatized and shunned by bosses, colleagues, family, and society, leading to loss of motivation to change and eventual relapse.

Some of these issues are discussed in detail in the following sections.

Stigmatization

The Yellow Ribbon Project,⁴ a community rehabilitation project targeted at helping ex-offenders reintegrate back into society, has done remarkably well in dealing with the more tangible issues ex-offenders face post-release such as finding a job. Many ex-inmates, however, still feel discriminated against by society due to their past prison record. Commonly described as “their second prison,” this refers to their psychological and social imprisonment upon release with the “keys” held by the ex-offender’s family, friends, neighbors, employers, colleagues, and the community at large. How much is this perception as opposed to reality? What is the extent of this stigmatization?

At Work

Work is an important component in the rehabilitative process for offenders. Gainful employment contributes to the lowering of the recidivism rate by boosting their self-

esteem by being able to earn a living to support their families. While tremendous support is provided by the government to facilitate job placement for pre-release inmates, many unfortunately still choose to view the opportunity as discriminatory due to the blue collar, entry-level nature of the positions despite predefined promotion prospects depending on performance.

This faulty starting assumption affected their attitude toward the employer even before work had commenced. Coupled with significant insecurity and self-esteem issues beneath their tough exterior, any criticism of their work by their boss or colleagues was often interpreted as bias. In one situation, an ex-inmate employed by a car washing company accused the supervisor of prejudice when he was not allowed to perform cash-related tasks. It was only upon other new employees joining the company that he realized all “newbies” were treated in the same manner. While undoubtedly there were actual instances of discrimination, the ex-offender was also often at fault for a wrongful attitude or poor work performance.

Low starting pay was another common gripe. Instead of trying to understand the cause, such as a possible lack of experience or qualification, the automatic yet questionable assumption was that it was due to discrimination in view of their prison record. In one extreme case, as illustrated in the negative branch diagram depicted in [Fig. 27-2](#), acceptance of a Singapore Government-assisted *Prepare and Place* (PNP) job and its accompanying low entry pay became the assumed root cause of every predicted future negative event in this inmate’s life. Instead of being hailed as a helping hand, the PNP job was ironically perceived as exploitation and the central potential cause of reoffense and failure.

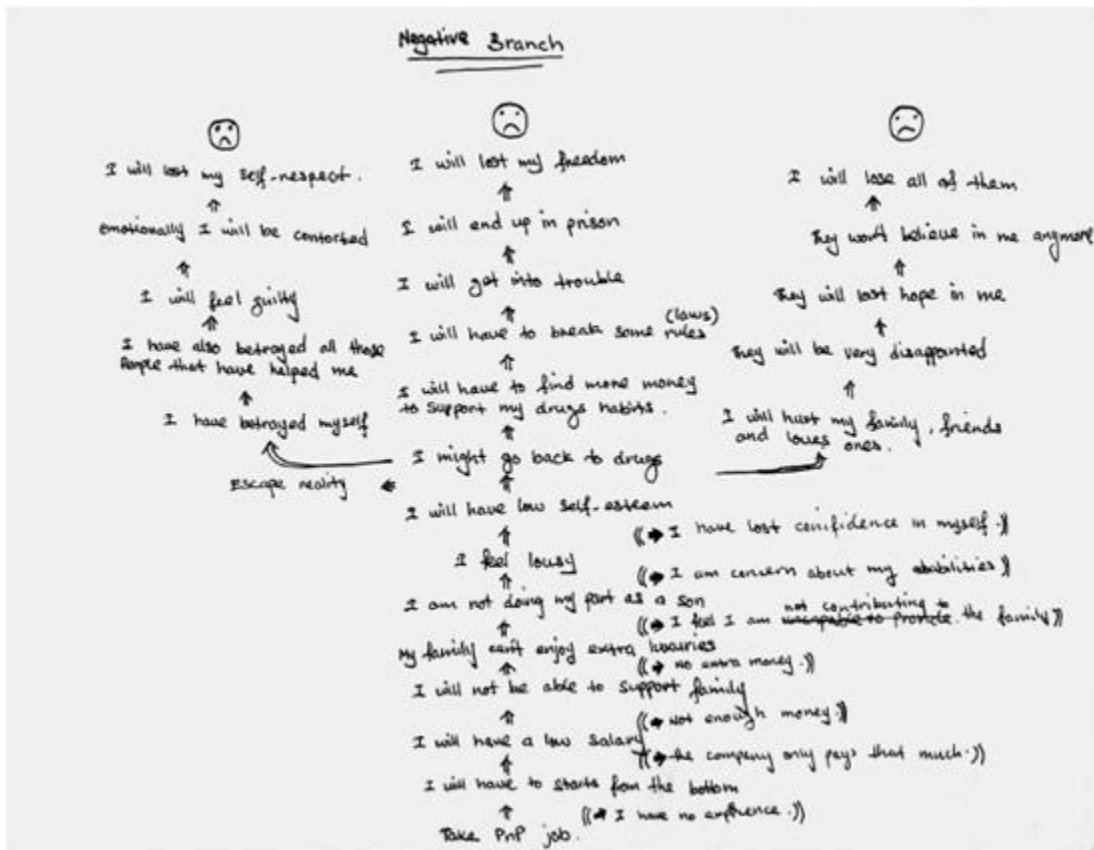


FIGURE 27-2 Negative Branch diagram.

At Home

Fear of stigmatization was by no means confined to the work environment. Many had a long history of estranged relationships with their family members because of their delinquent behavior. In almost every case, this was worsened by each side's tendency to see the worst, rather than the best, of the other due to unchecked and unchallenged erroneous assumptions formed purely from individual past experience, which did not allow for change. Even before release, many inmates were worried about the skepticism and perceived lack of support from family members in their attempts to start anew. Many complained about family members who would nag and call incessantly

during the day or even “spy” on them to ensure they were not hanging around with bad company, despite their genuine efforts to change.

For this reason, many found it difficult to face family members upon release. Even though they had nowhere to stay, many were afraid to go home until they could prove themselves or feel of value to the household. In one case, an inmate refused to stay with his sister upon release despite her pleas for fear that his brother-in-law would “look down” on him. After doing the TOC Cloud and identifying possible faulty assumptions about his family, he took the courage to face them with his fears and was transformed when reassured of their love and concern for him despite his starting halting offer: “I have nothing to offer. All I can give is just a kiss.” Two years later, this extended family unit is still intact.

Another reason for avoidance was the fear that family members would shun them or gossip about them. To avoid gossip, many tended to avoid family functions such as weddings and Chinese New Year celebrations, which creates huge inner conflict due to the importance of Asian filial piety, family ties, and kinship. Many secretly resumed illegal activities almost immediately upon release to earn extra money for the family in a desperate effort to prove their self-worth. Without the family’s knowledge or consent, one broke curfew during parole supervision to earn extra money while another accompanied his friend on a drug run to meet household bills. Paradoxically, upon learning about the illegal activity, their families refused to believe the best intentions of their behavior, leading to a deeper spiral of mistrust between both parties.

Other family conflicts remained deadlocked for years due to an inability to identify and resolve the core problem. In an example illustrated in [Fig. 27-3](#), an inmate was convinced that his mother hated him because she made no effort to reconcile with him for over 10 years. By simply reframing

and reverbaling his thoughts using the TOC Evaporating Cloud, he was stunned to realize that an alternative underlying need for his mother not wanting to reconcile with him may be due to “not wanting to be hurt again” rather than his longstanding dogged belief that she “hated him.” From the immediately softened expression on his face and subsequent positive actions taken to reconcile upon release, it was apparent that a once hopeless situation had given way to optimism. The power of just a few words cannot be underestimated.

Negative Peer Pressure

Negative peer pressure is cited as one of the biggest obstacles to successful rehabilitation. For many offenders, their negative peers represent their de facto family, or brotherhood, especially when they have come from families with little or no parental support or supervision. As a result, there is a disproportionate amount of loyalty and “stickiness” in these relationships, many of which are formed during the impressionable teenage years that provide a sense of belonging, security, and self-esteem. Unless there is an alternative way to satisfy these underlying needs (e.g., reconciliation with family, success in work, or finding a new circle of friends), these relationships are almost impossible to wean.

I have been in conflict between me and my mum since 1999, the year which I went to Reformative Training Centre for rioting. Actually my mum had already warn me when I was sentence to Boy’s Home in 1998. After I’ve been released with electronic tagging, I started to build a bridge toward my mum by going back to school but in the end, the bridge demolished. It goes same until now, when ever the bridge been build, it destroy by

me. So last year, I started to build it once more, it goes the same. My fault is, I never keep my promises to her....

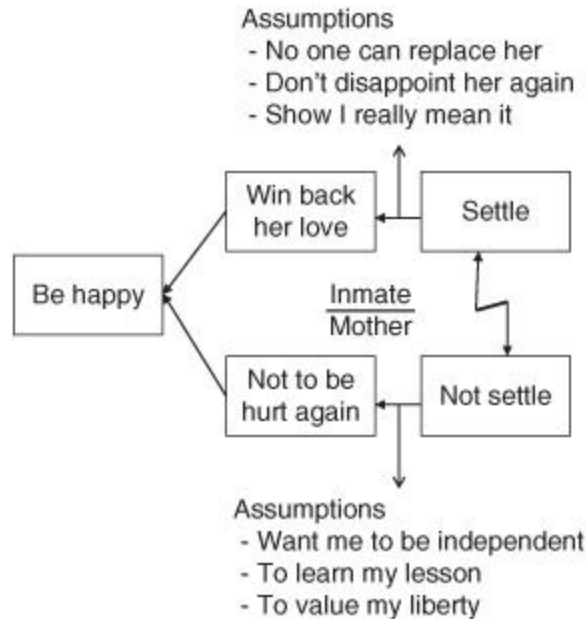


FIGURE 27-3 Cloud diagram.

On a very simplistic level, we can liken this to a child who, whenever he falls down, immediately cries for his mother. In the same manner, whenever the ex-inmate feels threatened in a stressful work environment, the tendency is to run toward the comfort zone of negative peers who provide both emotional and financial support. The core problem lies in the latter form of support rather than the former, which invariably in this group of peers is not easily delinked. Conditional to having this emotional security and acceptance, one needs to behave in a manner acceptable to the group. In more cases than not, this involves delinquent behavior to obtain “easy money,” which, in a very twisted logic loop, brings about an even greater sense of acceptance, achievement, and self-worth from their peers.

Consider the following highlighted excerpt from [Fig. 27-4](#) of a crudely constructed TOC *Current Reality Tree* (CRT)

skeleton drilling down from why the author does not see the need to change. The only question asked at each level is “why?” Even though the logical links are somewhat flawed and incomplete, it is astonishing to see the level of honest self-reflection after a brief two-hour exercise bearing in mind his lower education level and limited verbalization ability. From a starting position of blame (highlighted *undesirable effects* [UDEs]), the core problem shifts to self. Notice the importance of the branch originating from his desire to “feel famous and recognized” or, put another way, to have a sense of identity, which is provided from his peers.

Importance of Face

Linked inexorably to these issues is the Asian intangible concept of “face,” which is used in relation to honor, and its opposite, humiliation. Loss of face is linked to the fear that others may think badly of you, will not respect you, and will laugh and whisper about you behind your back. A similar term in Malay is “malu,” which means social shame, the inner feeling of doing something wrong and letting others down. In Asian society, protecting against loss of face becomes so central an issue that it swamps the importance of other tangible issues at stake.

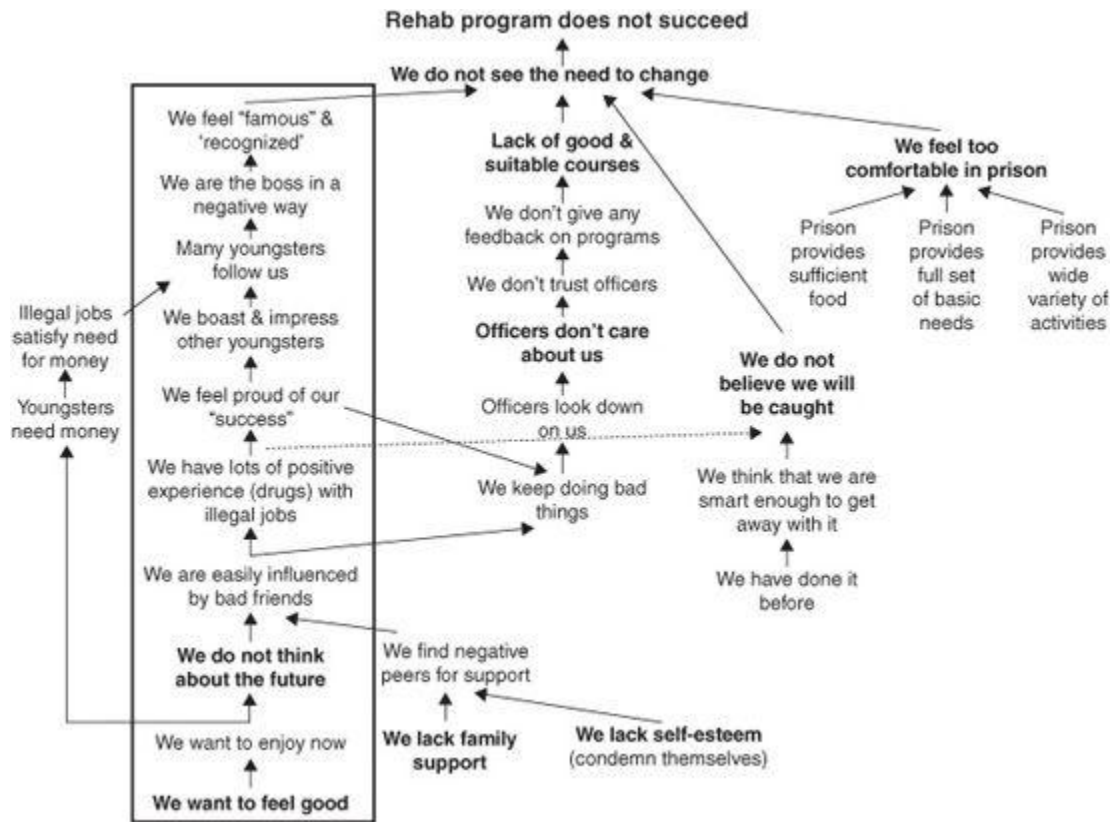


FIGURE 27-4 Current Reality Tree skeleton.

For inmates, the importance of face is even more pronounced due to their low self-esteem. With little to boast about except their negative achievements, the need to protect their honor and face becomes tantamount when threatened. Whether in prison, in the workplace, with friends or at home, humiliation is to be avoided at all costs, which often leads to seemingly irrational or illogical behavior. In order to save face, many feel that they have no choice but to make less than optimal decisions, which affects both their work and personal lives.

As an example, one inmate shared that he was “forced” to commit robbery because his best friend needed money for his mother’s hospitalization bill. At the time, he felt he had no other choice. The irony was that it was not at his friend’s request but his own need to feel “the man” and act as “big brother.” Without even considering other alternatives such

as government assistance for low-income families or consulting the hospital social worker, he automatically assumed the burden to meet his own misguided need to prove his loyalty.

In another incident, an inmate was arrested for vehicle theft and rioting after receiving an urgent call from a fellow gang member who was outnumbered and involved in a serious “showdown” with his enemies. With no money, he felt he had no choice but to steal a vehicle and go to save his friend. Based on this logic, he failed to see the justice behind his arrest and the reason for imprisonment was lost on him. In another rioting case, the inmate recounted that he had no choice but to fight because an enemy gang member had “stared at him.” In order to protect his honor, he had to stare back and fight.

From these incidents, one can see the extreme measures taken to save face. To them, choice is not an option where face is concerned. Unless they are able to look forward and question their priorities of what is really important in their lives, everyday actions will remain impeded by their unvalidated need to preserve face at all cost, even at the expense of work and family.

What to Change to?

Self-Regulation

Based on the previous observations, though by no means exhaustive, a pattern is starting to emerge. If we draw a conclusion strictly according to these findings, then presumably the main reason why ex-inmates do not successfully reintegrate into society is due to lack of control over external influences, that is, discriminatory employers,

lack of family support, negative peer pressure, prison life being too comfortable, and being forced by circumstances to commit the crime to protect their “face” and honor. If this hypothesis is correct, then it implies that the inmate is just a passive victim of circumstance.

The folly of this victim mentality is obvious. Undoubtedly, the role of the inmate supersedes any form of external influence. How we choose to think and what we choose to do is governed by self-will after deliberation on all external factors. The only way to improve our life is to take responsibility for our actions through self-regulation. As the old adage goes, *if we can't change others, then we can only change ourselves*. Using the TOC TP tools, the goal was to take them to the mirror and let them see their own reflection before deciding for themselves what makeover, if any, was required. Change must be prompted from within rather than dictated by others to be effective.

Why TOC?

Many people have asked what it was about TOC that led me to believe it could change deeply ingrained thinking patterns developed over a lifetime. How could we convince grown men of such tough demeanor to openly share personal problems, admit their personal shortcomings, and put them up for group scrutiny within an impossibly short contact time of 18 hours? What generic tool could meet the individual needs of 60 inmates with a plethora of different backgrounds and chronic conflicts?

To an observer, it seemed almost foolhardy for us to continue given our limited experience and intuition in the challenging prison environment. What was so special about TOC that gave us the confidence to carry on?

Three characteristics about TOC were pivotal to our decision:

1. TOC tools are Socratic.

From an early age, most inmates have been lectured on the right way to act, think, and behave by parents, teachers, social workers, counselors, and prison officers. As a result, like smokers, they have become numb to outside opinions, no matter how rational. TOC tools, however, provide them the freedom of choice to use their own words, expressions, and language to develop privately their own solutions to their problems. Once they understand the process, participants have free rein to express their own point of view within the parameters of a logic diagram without third-party interruption or distraction.

When these inmates, who are so used to not helping themselves, decide to buy in, then it is an enormous sense of accomplishment because they feel like they helped themselves. This empowerment facilitates ownership of the solution instead of having to be reminded constantly about the negative consequences of their actions. Within the setting of a correctional facility where inmates are largely constrained in the way that they act and behave, the importance of this sense of empowerment by unleashing their thinking processes cannot be underestimated. As one inmate put it, "You can control how I behave, but you cannot control my mind."

2. TOC tools cross all boundaries.

TOC tools are generic enough to use across different industries and applications but yet specific enough to meet each participant's needs irrespective of age, education, and culture. They provide a simple yet logical framework to check one's thought processes in the language and vocabulary with which one is familiar and comfortable.

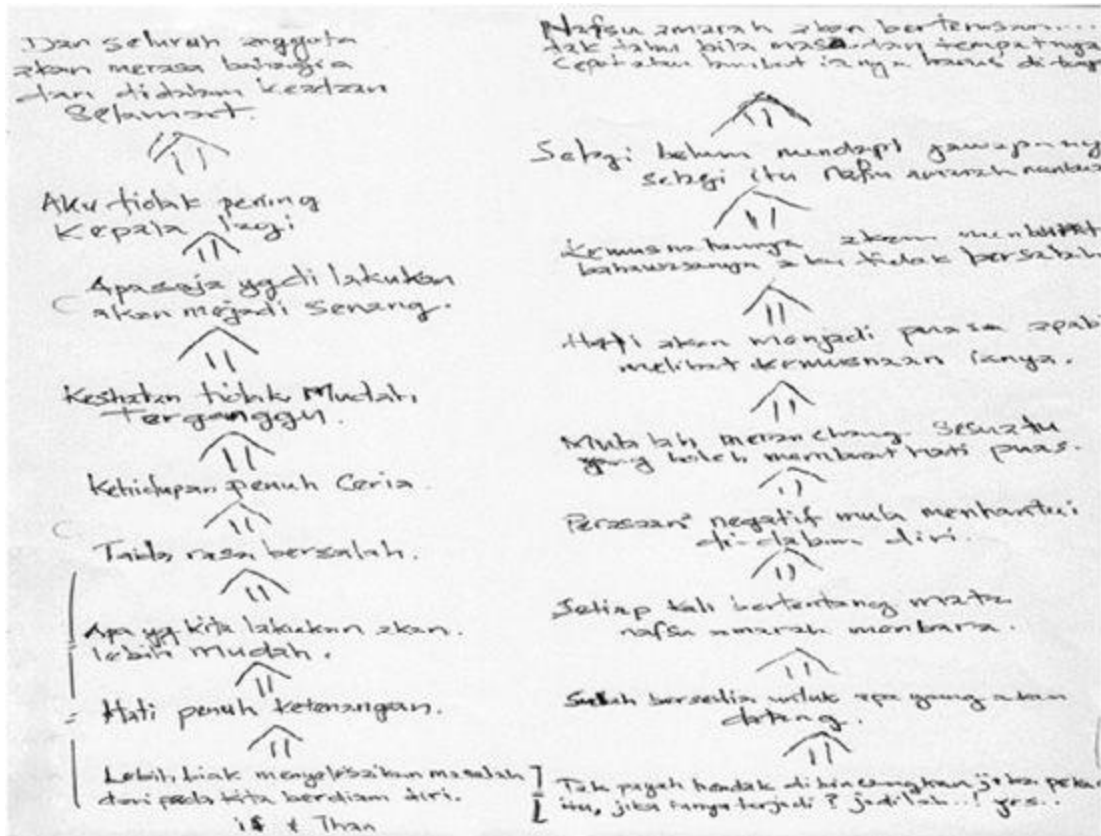


FIGURE 27-6 Branch diagram in Malay.

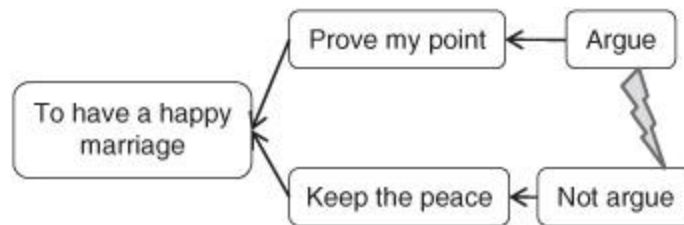
By using the TOC Evaporating Cloud, one is able to logically validate one's understanding of each situation and identify the erroneous assumption that has led to the wrongful action. Coupled with the self-correcting terminology of the tool, one learns to see the best in self and others through identifying the positive need underlying each action and thus avoiding preemptive and wrongful judgment.

Figures 27-7a and b illustrate a marital conflict. In order to have a happy marriage (common objective), I must prove my point (need). In order to prove my point (need), I must argue (want). Does proving a point lead to a happy marriage? Is arguing the only way to prove a point? We can rewrite this cloud using different wording: In order to have a

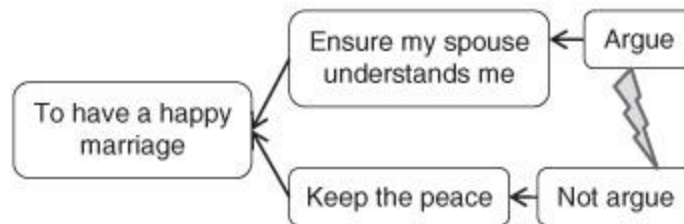
happy marriage (common objective), I must ensure my spouse understands me (need). In order to ensure my spouse understands me (need), I must argue (want).

The difference in wording between “prove a point” and “ensure that my spouse understands me” is slight but the difference in meaning is enormous. The first is all about self, but the latter implies that the spouse’s opinion matters. Other ways to increase understanding could be to write a letter or e-mail, or have a quiet chat.

By learning to stand in other people’s shoes, one becomes more empathetic and receptive to new and different points of view and opinions over time. For inmates, this is an excellent reflective tool to review the actions of others, as well as self, rather than just believing in and accepting the worst of self and others.



a. Inmate's initial description.



b. Inmate's revised description.

FIGURE 27-7a and b Marriage conflict.

How to Effect the Change?

Armed with a portfolio of TOC TP tools, we set to work in planning the curriculum. Time was the major constraint at this juncture as we had exactly one month to complete our due diligence and develop the course materials. At the same time, we were faced with the problem of how to train such a diverse audience. While the tools themselves were generic enough to address individual needs, we did not have the luxury of working one-on-one so we needed to redevelop the training materials to address all learning capabilities within a classroom environment. Anyone who has worked with such a heterogeneous audience will understand the enormity of this task. The main areas to address were marketing, course materials, and delivery.

Marketing

Marketing can be described as *the activities required to persuade a client to perform a certain act or transaction through identifying and meeting the client's needs*. In this case, even though the pilot study had already been approved by Singapore Prison Services for implementation, it was equally, if not more, critical to buy-in the end client, namely the inmate audience. To achieve this buy-in and ensure a high level of understanding, heavy customization of our TOC materials was required to meet their needs. Details of when, why, what, and how the marketing was done is detailed in the next section.

Up Front Buy-in

Successful buy-in of the inmate audience was critical to the success of this project. Typical TOC trainings commence with a brief marketing overview followed by a presentation of process skills. This approach is fine when the audience already recognizes the value of TOC, for example AGI Jonah

or Goldratt School participants who have opted to invest significant time and money on a course and are further along the buy-in process. As a complete contrast, our inmate audience knew nothing about TOC except that it would help them escape from the monotony of their cells.

For such a mandated audience, a strong buy-in as opposed to a marketing overview is crucial to ensure the right attitude toward the course is adopted. In a typical social work setting, it takes many sessions over a period of months to establish trust and rapport, which is not always a given either. Unless the trainer has something that the clients want or they feel the trainer can help them with, they will not buy in. In business, a similar problem but to a lesser degree is encountered with staff attending trainings to merely fulfill training hours or management instruction. Unless they truly value or buy-in to the subject matter, it is not a sustainable process. Adequate course time must be allocated for buy-in, which unless performed successfully up front, impedes on the depth of the process being understood and taught.

While it may be argued that the buy-in can be achieved progressively throughout the course, our experience has shown that it is more effective at the onset as TOC is a process that requires concepts to be understood in a sequential manner. Even for the “converted,” TOC, with its own set of rules and vocabulary, is often regarded as a tedious process, and requires a certain level of stamina to master. Layer One of the TOC Layers of Resistance (Identification of the Problem, see [Chapter 20](#)) must be established early, otherwise a serious compromise is made between the quality and quantity of material learned.

Motivation for Buy-in

How do we motivate an audience to buy-in to TOC? The marketing buy-in for TOC behavioral applications is vastly different from that of business applications. Unlike business applications, where management is motivated to implement TOC for higher profits and employees are obligated to follow instructions as part of their job, it is far more difficult to convince another individual, in the absence of any immediate tangible benefit, to change a behavior that has been developed over a lifetime. What could motivate someone to change the way he thinks, behaves, and reacts, especially within such a short time period?

The goal of TOC is to challenge our way of thinking, behaving, and making decisions. The paradox lies in the logic that in order to motivate change, one needs to *prove* that there is a need to change. For a thinking skills program, this implies that existing thinking is flawed or suboptimal. (Anyone who has tried to correct their spouse, even with the noble motivation of improving their relationship, will empathize with this task!) Without accepting or understanding that there is a need to change, this creates resistance, fear, and distrust. Before we can market something new, we need to prove that their existing approach may not be optimal and complete this task in a nonconfrontational and nonthreatening manner.

The Buy-in Process

The marketing process contained five steps:

1. *Communication*. The first step of the marketing process was to find a way to communicate with our target audience in a manner that they could easily relate to and understand. As buy-in was critical to program success, we needed to ensure our message was clear and relevant to their needs. In addition to the internal focus group sessions conducted with

officers and inmates within Singapore Prison Services, we consulted with a number of ex-inmates and their families, counselors, employers, prison fellowship church groups, and charitable organizations involved with prison rehabilitation to gain a wider and deeper understanding of their personal, home, and work environment viewed from the eyes of different interest groups. Again, one of the most effective ways to gain considerable insight was through the simplified use of the TOC PRT framework of identifying obstacles through informal conversation.

2. *Customization.* The next task was to customize the buy-in in relation to overcoming constraints in the workplace because our ultimate measure of success was job retention upon release. The significance of the workshop title, Reintegration, with regard to both family and society, was selected to engage their interest as pre-release inmates. Although it was obvious from our research that it needed to be a personal paradigm shift, refocusing the course title on work rather than self allowed them to concentrate on the process rather than worry about how they would be viewed by others. Conceptually, the learning process was the same but it would be less intimidating and raise less negativity during the buy-in process.

3. *Validation.* In the course of buy-in, a number of directed activities and exercises were designed to challenge their thinking by disproving their logic in a nonthreatening manner. Linking the activities together was the ongoing need to ask the underlying question "Why?" which is also the lowest denominator of the TOC TP. As to be expected, the practice of questioning every action in daily prison life is not high on the list of skills encouraged in a correctional facility, and to

reawaken this questioning ability after years of incarceration was like trying to crank a car engine that had been left in the garage for years. Once started, however, it was raring to go and difficult to switch off again.

Indirectly, the aim was to prompt and question the logic and clarity of one's own thinking process and belief system. Learning to question self, though, can lead to harsh realities. The key was to downplay the underlying self-directed activities under a common reintegration theme, which eventually led to an open discussion as to the validity of their own thoughts, words, and actions in a reflective and yet fun-filled and collaborative manner. The training bonus was the resultant vivid transformation of a group of wary individuals to a bonded team who could openly laugh at themselves and at each other without a sense of embarrassment or loss of face.

4. *Secure Environment.* By creating a secure and safe learning environment for the inmate audience, we could encourage and maintain open dialog without fear of mockery, judgment, or reprimand within a closed confidential circle. For this same reason, we decided to conduct the course without the presence of any "uniforms" or prison staff in the classroom. (The obvious concern for the trainers was a safety issue but our fears proved unfounded after being well equipped with shrill alarms and constant closed circuit television surveillance monitoring by the prison officers.)

5. *Time Allocation.* Of the entire 18 hours allocated for the workshop, approximately one-third or six hours of the content went into the buy-in process. The remainder of the time went toward teaching three selected TOC TP tools—the Evaporating Cloud (EC),

Logic Branch, and Prerequisite Tree. From a planning perspective, it was agreed from the onset that the number of tools taught would be sacrificed for more buy-in time if, as, and when necessary. We strongly felt it was more important that the audience, many from a low educational background, left with strong foundational skills rather than a sketchy recollection of three tool processes. Fortunately, this compromise never took place. Conversely, the more time that we spent on buy-in, the faster the teaching of the tools went. For most participants, this first 6 hours proved to be the most invaluable part of the course.

Why is it so hard to answer the question why?

Ironically, as many TOC for Education Inc. practitioners will agree, it is much easier to teach the same TOC tool to a child than to an adult. The purity of the answer to the question “why” seems to diminish in proportion to age. Likewise, it is the same when training a group of lower level employees versus senior managers. Why should that be? After much deliberation on this topic, I can only concur that our minds become so overloaded with information that it becomes harder and harder to extract the core essence of our thoughts, which TOC tools help us do so perfectly. Unlike children, our words are so intertwined with political correctness and societal expectations that eventually over time the true meaning is no longer communicated.

Course Materials

From a course developer and trainer’s perspective, the end goal was to ensure that the learning was sustainable. In the absence of any immediate tangible benefits to the user, the problem was not so much in teaching TOC processes given the high logic component in Singapore education, but how

to ensure continued behavioral application under stressful situations when the need for TOC is the greatest yet automatic default behavior takes over in the flight for safety and security.

Over the years, a common observation from local TOC courses conducted in a variety of corporate, school, and social services settings by experienced TOC trainers was the relatively low adoption rate after the course. While a certain percentage of buy-out is expected after training, it was surprising that a high number of those participants who seemed clearly bought in during the course rarely applied TOC afterward. Despite glowing feedback reviews and excellent process skills, few participants seemed to apply the tools on a regular basis after the course.

After contacting a number of past course participants, many admitted they enjoyed the course but felt that the TOC processes were too tedious to repeat. The compulsory use of TOC terminology to derive the logic such as “in order to . . . I must . . . because. . . .” and “if . . . then . . . because” and its rigorous process steps were considered too time consuming for regular use. While most felt the concepts and processes were relatively easy to grasp, they were not prepared to invest time and effort to practice and use TOC for daily issues that they believed could be solved intuitively without the need for any special thinking tool.

For others, they found it difficult to find appropriate daily opportunities to practice the tools. Not every decision they faced required a logic branch nor represented a fullblown conflict and not every conflict required a resolution. While some may argue in TOC circles that everyone has conflicts, it also depends on the severity and the outlook of the individual for the tools to be relevant. As a result, even when participants bought-in to the concept and were able to apply it effectively to case studies provided during the course, the knowledge faded after the course either due to a perceived lack of opportunities to practice or a

subsequent buy-out after finding the tool too time consuming to use for daily problems.

Given these existing constraints facing our “educated audiences,” it was critical that we find a way to prevent “mental indigestion” for our inmate audience who were likely to find the processes even more difficult to use and practice within the narrow confines of prison life. Two fundamental questions needed to be answered, namely:

1. Was it possible to distill the TOC tools into their core components to simplify the learning process further?
2. How could we expand the opportunities for TOC participants to practice and use the tools?

With these questions in mind, we needed to create simple TOC materials applicable across age, education level, and language with easy applications to daily life.

Core Content

The TOC TP Tools are based on two types of logic—necessity and sufficiency, and the concept of win-win. Rather than launch directly into the mechanics of the tools, however, for the reasons mentioned previously, we decided to teach these principles first to simplify the learning process. Once these principles were well understood, the ability to apply the tools would subsequently fall into place.

Teaching Necessity Logic Both the EC and PRT tools are founded on necessity logic. Both are read in an “In order to....we must...because...” format and the validity of their cause-effect relationships depends on meeting minimum necessary requirements. In many instances, we saw amazing breakthroughs through the use of these tools, but their full application is intended for more significant issues rather than to overcome daily run-of-the-mill decisions (e.g.,

choosing between buying apples or bananas), although the necessity logic underlying both types of applications is the same.

The underlying logic is straightforward. Every action that we take is driven by an underlying need. As shown in [Fig. 27-8](#), in order to make logical decisions, we need to:

1. Question the validity of that need,
2. Check whether there is a better way to meet that need,
3. Check the underlying assumptions if necessary.

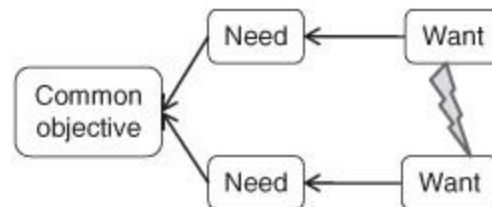


FIGURE 27-8 Cloud with common objective, needs, and wants.

Differentiating between Needs and Wants To teach necessity logic, we focused on the main component of the TOC Evaporating Cloud—the relationship between the “need” and “want” on either side of the conflict. The first step was to teach the importance of needs over wants using many different group activities such as basic budgeting, needs analysis, or demonstration games like the Potato Experiment (see shaded box). Considerable time was spent on this topic, as understanding this concept was central to learning subsequent TOC TP tools.

Props Required—One clear plastic container, one bag of mixed potatoes ranging from large to baby potatoes, one bag of uncooked rice, and a 1-L bottle of water. Show only the plastic container and potatoes.

1. Ask a volunteer from the audience to fit as many potatoes into the container as possible, preferably in order of largest to smallest. They will soon realize that the trick is to insert the largest potatoes first and then intersperse the gaps with the smaller potatoes. Now ask the class whether the container is full to which the answer will be yes.
2. Reveal the bag of rice and ask another volunteer to pour as much rice as possible into the container. Once done, ask the class whether the container is full to which again the answer will be yes.
3. Finally, reveal the 1-L bottle of water and ask another volunteer to pour the water to fill the remainder of the container. Once again, ask the class whether the container is full to which the answer will be yes.
4. Let the container stand for the remainder of the training session. Over time, you will observe the water slowly being absorbed by the rice grains, which become enlarged and in turn force the top layer of potatoes to slowly pop out of the container. (Point out to the class how fast and easy it is to pour the rice and water into the container compared to the difficulty in fitting in the potatoes. Once everything is added, it is very difficult to remove the wet rice, which sticks to the potatoes and container, and even more impossible to remove the water upon absorption.)

5. Ask the class what they learned from this exercise assuming that:

Container = our life

Potatoes = our needs (sized in order of importance)

Rice and Water = our wants

Lesson of the Story

Our lives, like the container, have a limited capacity so we have to choose carefully what to include. Even when we start with a very clear focus and sense of priority as to what are our “biggest potatoes” or most important needs in our life such as love, family, freedom, we are often distracted by the “rice” and “water” or non-essential wants such as pride, popularity, and easy money which can take over our lives before we realize it.

As a result, we must constantly prioritize what we choose to put into our lives and protect our underlying needs. Over time if we are not careful, our needs can easily become dislodged by our wants, which grow in false importance and impact on our thoughts, words, and actions. We need to clearly define our needs and make sure they are well grounded in our life to prevent being overshadowed by non-essential yet competing wants.

© TOC ASIA PTE LTD. All Rights reserved.

Identifying Underlying Needs Once the audience was able to differentiate clearly between the concepts of wants and needs, the next step was to enable them to understand the relationship between a want and its underlying need by asking the question “why?” While this step was relatively

easy for our inmate audience, who had already internalized the practice of questioning and asking “why” during the first few sessions of buy-in, the main difference now was teaching them to *evaluate the logic* of their answer by using the TOC terminology and framework “In order to . . . I must . . .”

A typical illustration is a smoker’s desire for a cigarette (want) whenever he needs to relax (need). If the need for relaxation is validated, the next question is whether there is any other way to satisfy that need. Unless he can find another way to satisfy this need through other relaxation techniques (for example, to exercise, chew gum, or listen to music), smoking remains his default action. When we opt for one action over another, the implication is that this is the *only* way to meet this need at that point of time; that is, *In order to relax, I must smoke*.

To many, the absoluteness of this wording is difficult to accept. The usual response will be a retraction or a disclaimer that no better options were available at that time. Once again, we need to question whether this is true. The irony is that while we have the freedom of choice to exercise actions that are more appropriate, we simply choose that which is familiar. By default, we do not question our actions because our responses have become automatic after years of practice. From our experience, this is a sure way to meet our need, regardless of whether it is the optimal action to take.

In the classroom, there are numerous opportunities for the trainer to let the audience practice this skill. Focusing on actions, enables the audience to practice necessity logic in a far wider range of situations than conflicts. Most people only have a limited number of conflicts at any one time, while desired actions are plentiful and easy to identify (e.g., why buy a new hand phone, why eat an extra doughnut, or why go on vacation). Even if nothing else is learned for the remainder of the course, the benefits of learning to question

one's actions by challenging one's belief system before acting is immeasurable.

Validating the Need One of the most powerful exercises that we conducted was to ask the audience to write down their crime (want) and then ask themselves "why" (need) they committed that crime. Typical responses were:

NEED (In order to . . .)	WANT (I Must. . .)
Have money	Steal, traffic drugs, rob
Have respect/prove myself/impress others	Riot (fight)
Feel good/relieve stress/enjoy	Consume drugs
Be accepted	Join a gang

At first, their answers were assured and confident. However, upon deeper questioning as to the validity of their need using the TOC questioning framework "in order to...I must," the surety of the answers began to waiver. In almost all cases where money was the stated need, it turned out to be not for real financial woes but to address self-esteem needs such as to show off their wealth, to prove their loyalty to gangs, to impress others by being "the man," as well as for the Rolls Royce of all needs—to obtain "easy money" or put more bluntly—to avoid hard work.

In order to obtain easy money—I must commit crime

Now behind bars, the rhetorical question for inmates became whether in fact, because of their crime, they met their need to show off or enjoy their easy money. Sitting in a circle on the hard concrete floor, the group conclusion was that perhaps easy money was not so easy after all!

Others, who committed a crime to fill their need to impress their friends, were suitably chastened when they admitted that their negative peers, whom they sought most to impress, had completely disappeared after their arrest. Instead of admiration from their peers, the result was

avoidance. Only their family remained to support them through their incarceration.

In order to impress my friends—I must commit a crime

This invalidated their original need by teaching the hard lesson that their need to impress may have been directed to the wrong party, leading to the wrong action.

Finding an Alternative Way to Meet the Need If the need is validated, we need to question whether there is any other way to fulfill that need before taking action to achieve the desired want. The objective is to open their minds to different possibilities to fulfill their need. An illustration of one group's solutions to fulfill their ongoing problematic need for money and how to find new (positive) friends is shown in Figs. 27-9a and 27-9b.

Many of our audience were in prison for drug-related crime. Even though they had been “clean” during their entire incarceration, many knew their own weaknesses and worried about their high probability of relapse upon release. When questioned about their need to take drugs, common answers were to feel good or high, to relieve stress, or to experience the adrenalin rush.

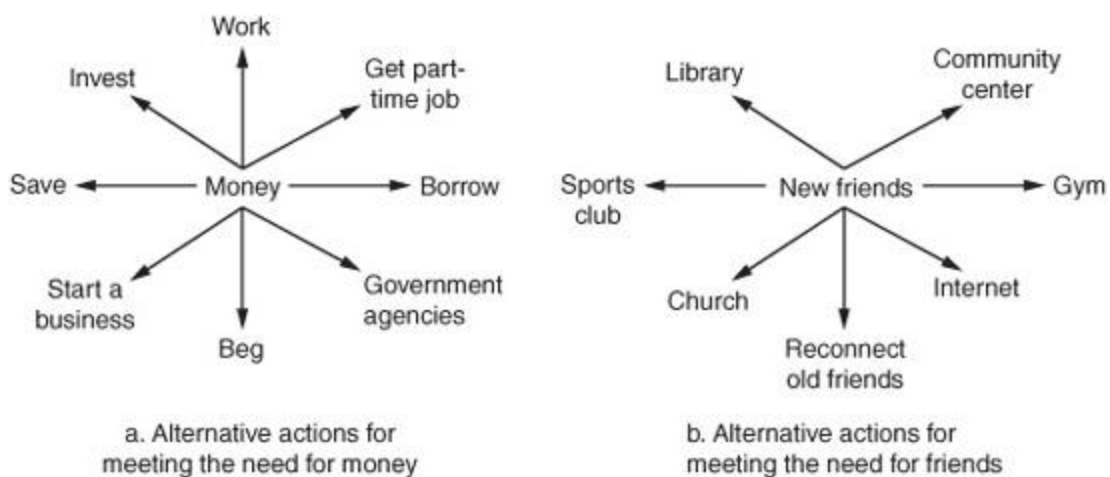


FIGURE 27-9a and b Meeting needs of the Cloud by alternative means.

In order to feel high—I must take drugs

Was there another way to feel high? One inmate excitedly put up his hand and suggested running to get the same adrenalin rush. Apparently, he had been a school athlete and loved to run.

In order to feel the adrenalin high—I must run

Did it work? Upon release, he contacted us several times to inform us that he was still clean and still running! While this was obviously not the solution for all, it emphasized the importance of finding an alternative way to meet the need; otherwise, the tendency is always to go back to default behavior.

In a case of theft, one inmate shared that he stole a \$30,000 luxury watch after trying it on in the shop and admiring it on his wrist. His need for stealing the watch was to look cool and impress his girlfriend.

In order to look cool—I must steal the watch

Was there another way to look cool? Standing there in his decidedly uncool prison uniform and rubber slippers while sharing his story, it suddenly dawned on him that there were other ways, such as changing his hairstyle or becoming a good dancer, that he could have looked cool and impressed his girlfriend.

Practice examples were by no means confined to reflection of past experiences. Everyday actions were perfect to practice necessity logic. On one occasion when TOC course time happened to clash with exercise yard time, one inmate angrily rushed up to a prison officer and threw his file on the floor. Fortunately, no charges were made but

when questioned why he behaved in such an aggressive manner, he explained that he wanted to get the officer's attention.

In order to get the officer's attention—I must throw my file on the ground

When asked if there was any other way to get attention, he sheepishly shuffled his feet and murmured that he could have waited until the officer was free. Although the rest of the course still clashed with yard time, he became a model student.

In another memorable case, a younger inmate who was due for release the next day had his sentence extended for tattooing his forehead with a pencil. Upon returning to class, we questioned his need to tattoo his forehead, leading to his reply that he and his cellmates were so bored the night before that they did it for fun. The class collapsed in hysterics when he read out from the board, *"In order to have fun, I must tattoo my head."* Was there no other way to have fun? Today as a free man, when he calls to chat, he reminds me of his hard lesson learned in his quest "to have fun."

Learning to question oneself can lead to harsh realities, making transformative experiences painful to go through. Unless there is critical reflection time, there is no point in attempting to teach the additional process and language associated with the rest of the tool until fundamental principles are understood. Usually one-on-one intervention is required for chronic behavior modification; however, the safety perimeters established within this group were so tight that they were able to lower their tough façade and openly share and see the folly of their behavior from another pair of eyes. A common phrase heard in the classroom was an incredulous, "You did what? Why?"

By breaking down the Cloud process, it also allowed the audience the opportunity to analyze the "why" behind their

actions minus the usual accompanying “why not” during the initial learning process. For most adults, our inherent sense of right and wrong is so strong that for most types of negative behavior, the “why not” is already well understood. For this audience in particular, well-meaning family, friends, teachers, and counselors had drilled them with “why not’s” since they were young. Like smokers, however, even though they understand the consequences of their actions, few could quit until they could find another way to meet their need.

These are but a small sample of the numerous simple yet transformational situations resulting from asking the basic question “why?” and identifying the core reason behind our individual actions. While it may be argued that these concepts are not unique to TOC, the use of the key TOC phrase “In order to . . . I must” was critical for success. The combination of the three (the need, the want, and the justification) provided a simple yet effective primer for the rest of the Cloud tool, which is essential to deal with interpersonal issues and more significant personal dilemmas that require a more thorough analysis of both sides of the conflict.

Win-Win Buy-in to the concept of win-win is essential to teaching the remaining Cloud tool process. The need for win-win, however, is not an easy concept to sell in the Asian context where it is not common to insist on one’s point of view, especially with respect to superiors, elders, or authority. Unlike individualistic societies that thrive on debate and the fundamental belief in the freedom of expression, traditional Asian culture does not encourage direct confrontation because being rebuffed could cause loss of face for either party.

Rather than striving to achieve a win-win situation, it is far more common and acceptable to adopt a strategy to avoid, give-in, comply, or compromise even if it means a win-lose

and ultimately lose-lose situation for both sides at the end of the day. To overcome this, once again we needed to *disprove their underlying logic* before trying to introduce new concepts. To give a sense of fun to the exercise, we engaged in active role-play to demonstrate the outcome of each type of conflict resolution.

To encourage ownership of the solution, the class made a list of all of their existing ways to resolve conflict before discussing the merits and disadvantages of each. Not surprisingly, the most common solutions were avoid, comply, give-in, and compromise. The clincher was in the realistic demonstration of each scenario by self-professed “actors” in the audience, which added to the high recall ability of the exercise for many months afterward.

Examples of role play include: a man agreeing to marry his girlfriend after she threatens to leave (give-in), a prison officer insisting certain actions to be followed by an inmate (comply), a mother nagging her son until he decides to move out (avoid), a husband and wife agreeing to take turns watching TV for 15 minutes each during the final match of the Soccer World Cup and a TV drama finale (compromise). By the end of this exercise, participants are much more receptive to learning about win-win upon seeing the consequence of win-lose.

Teaching Sufficiency Logic Sufficiency logic is the basis of the Logic Branch used to understand consequences of actions and improve half-baked ideas. It is read in an “if . . . then . . . because” or “if . . . and if . . . then . . . because” format to describe why situations exist or why we believe particular actions will result in certain outcomes. The validity of their cause-effect relationships depends on *sufficiency*.

The concept of cause and effect is well understood by most people. With very few exceptions, every inmate knew the immediate negative consequence of their actions prior

to committing the crime and yet still went ahead. What prompted them to act in such an illogical manner? There are two main reasons behind this:

1. Failure to understand the full consequence of actions
2. Failure to validate the predicted effect

Understanding the Full Consequence of Actions In behavior, *necessity logic* is critical to understanding *what causes us to act*, whereas *sufficiency logic* helps to validate what we believe *will happen* because of the act. The problem behind the latter is that it is often determined by our individual experience and intuition rather than from the possession of full facts. If we have insufficient intuition about the situation, then we rely on our limited experience to form an opinion. Based on these opinions, we form behavior patterns that govern how we behave and think.

For example, *IF I offend, AND I get arrested, THEN I go to prison.*

IF I offend AND I do not get arrested, THEN I make easy money.

Without sufficient intuition, the decision whether to offend is made based on individual assumptions about (1) prison, and (2) the probability of arrest. For first-time offenders, both sets of assumptions are based on the experience of their peers, which is often exaggerated to show off their bravado. For multiple offenders, it depends on their personal experience. In both cases, the intuition is usually inadequate to make a well-informed judgment within this limited circle of knowledge and the branch is ended prematurely. Guided facilitation is required to extend the logic branch into a deeper understanding of consequences; for example, easy money may result in negative influence of peers, loss of work ethic, and greater tolerance and propensity for crime.

The other problem associated with insufficient intuition is the inability to see the full consequence of actions beyond themselves that occur because of their offense, especially if there is not a clear victim as there is in a rape or murder case. While the negative consequence to self and their immediate family is clear, many feel exonerated for their offense after being charged and incarcerated, i.e., their punishment has already paid for the crime. What they do not realize is the domino effect of their crime to society and everyone within their sphere of influence.

[Figure 27-10](#) illustrates a drug trafficking example where one can see the initial logic branch resulting in either “I suffer in prison” on the left or “I get rich and impress peers” on the right. After much debate, the class was shocked to see the far-reaching effects on their peers, families, and clients as well as all those they directly influence. Consider the impact of one person who is trafficking drugs to 50 clients who in turn become addicted and in turn influence their peers, leading to a never-ending negative reinforcing loop. The impact of one person on society is enormous. Never before had they considered those who they did not know or could not see.

This butterfly effect⁵ perfectly illustrates how a single action can be magnified into an unstoppable chain of events and stresses the importance of understanding the full consequences of our action before we act. To cement their learning, each worked on their own crime cases through facilitated questioning, for a sobering analysis as to the implications of their actions on society. While it is impossible to predict all possible effects from each action, the aim is to create a heightened awareness of the implications to others rather than just self.

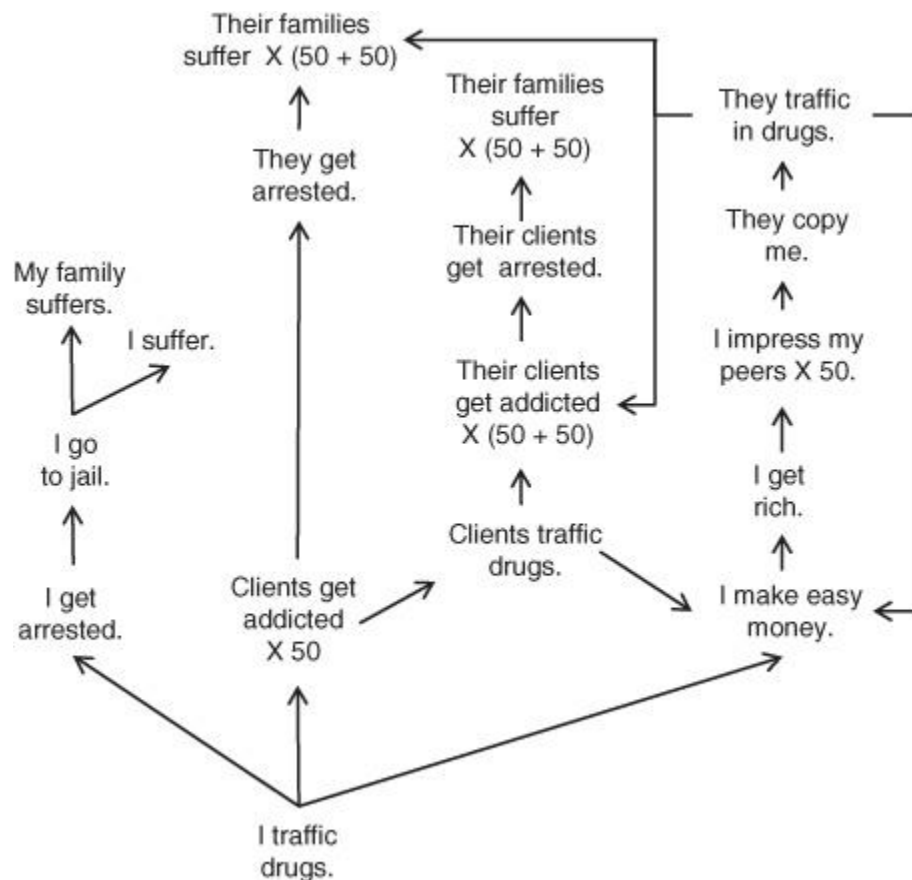


FIGURE 27-10 Consequences of drug trafficking on others.

Validating the Predicted Effect Apart from not realizing the full consequences of actions, many inmates fail to validate their predicted effect. Instead of bothering to check the logic of the predicted effect, many just blindly follow the belief system and behavior pattern of others. For example, IF you only consume drugs once, THEN you will not get addicted, or IF someone stares at you, THEN you must stare back. For others, even with knowledge and desire to the contrary, the default need to follow established behavioral norms takes over.

A common issue among inmates is whether to disclose their prison record to prospective employers. Based on their limited experience and hearsay, most employers will not hire them if they know about their prison record. As a result,

they feel they have no choice except to lie in order to get the job.

IF I lie about my record, **THEN** I will get the job, **IF** I get the job, **THEN** I will work hard to prove myself,

IF I work hard to prove myself, **THEN** the employer will retain me if he finds out about my past.

IF I do not lie about my record, **THEN** I will not get the job.

There are several glaring mistakes in the logic. Is it true that the employer will hire you just because you do not disclose your past? Is the employer really biased against inmates? Do you meet all other qualifications? Will the employer retain you once they find out about your past? In many cases, especially with service industries such as hotels, it is against company policy to hire an ex-offender. Separately, there are many other consequences of lying, which have not been addressed. An example of redefined logic after validating the predicted effect is shown in [Fig. 27-11](#).



FIGURE 27-11 Predicted effects of lying to an employer.

Like necessity logic, it is important to teach the concept before teaching the full tool. Using one's desired action as a cause, participants are made to practice single step branches using "If . . . then . . . because . . ." until they perfect sufficiency logic and are ready to draw the entire branch.

Delivery

The next challenge after simplifying the TOC TP into core components was how to deliver the content in a manner that they could easily process and understand.

Teaching Techniques

A variety of teaching techniques was instrumental to retain learning among inmates. Traditional classroom teaching was

impossible given the large variation in language and literacy levels among the group and made it difficult for trainers to engage all inmates at the same level and pace. Looking at the blank faces at the commencement of each course, it was often difficult to ascertain whether it was non-buy-in or just underdeveloped brainpower. To engage all levels, use of high-energy games, group work, individual reflection, and video presentations, each specifically customized for their unique lifestyle, helped to generate a high level of interest and maintain motivation and attentiveness.

The pictorial nature of the tools also provided a different learning dimension for inmates with different learning styles. In one instance, several TOC cellmates who had difficulty understanding the process of the Conflict Cloud were taught at night in their cell by another cellmate who was not a TOC participant, but by the nature of his occupation as a tattoo artist, could immediately understand and interpret the simple flowchart or pictorial nature of the tool. Others related to the tools as a form of challenge or puzzle, for example, a crossword or Sudoku, which had the added advantage of being able to help them work out their life issues.

Most importantly, the Socratic approach toward teaching TOC was a refreshing change for an audience who was so used to being told what to do. By helping them to find their “voice,” inmates gradually became more receptive and motivated about learning and applying the tools learned. For maximum recall, fun quizzes, worksheets, and notes were given to inmates at relevant junctures throughout the course. Inmates were also given “homework” to bring back to their cells, allowing them more time for individual reflection and informal group discussions while allowing trainers to effectively focus the classroom training period on delivering material.

Language

By default, the entire workshop was conducted in English due to language limitations on the part of the trainers. As a result, considerable modification to training materials was required to ensure the audience could follow the TP. In almost all cases, the target group could understand and speak simple English intermingled with local dialect, but the learning process was often hampered by weak vocabulary and communication skills. Much of the TOC terminology proved incomprehensible to the audience, resulting in heavy editing of original training materials as well as ongoing translation by self-volunteered translators within the group.

Ironically, even in their own language, many were often at a loss for words through lack of practice because everyday prison lingo tended to be abbreviated and colloquial, which was inadequate to express what they really meant. With this handicap, writing was an even bigger problem as evidenced by the tortured yet comical facial expressions of the inmates while complaining about their “brain jam” and “brain freeze” when asked to express what they felt at the end of each session.

In a strange twist, this constraint turned out to be a blessing in disguise as it resulted in enormous camaraderie within the group. Much collaborative effort was spent after class in their cells debating how best to define accurately their personal problems as evidenced in the high quality of homework. Forced to summarize their life stories within the small boxes in each TOC tool, they were taught conciseness and clarity of thought. On the part of the translators, their new role gave them a sense of importance and responsibility, while higher-level learning was reinforced through continuous internalization, interpretation, and repetition. In true TOC fashion, the burden of individual conflicts soon became shared group concerns.

In hindsight, the forced slower delivery of the more difficult parts of the course allowed the audience much

more time than average to think and reflect. Most importantly, it forced us to condense the materials into the simplest denominator for basic understanding, which drove us to the core of TOC.

Duration

Attention span was earmarked as a potential problem from the onset as most participants had not stepped into a classroom environment for more than 10 years. Not unlike young children, many participants initially found it difficult to stay still and focused for long periods, so we needed to provide constant group activities and breakout sessions to retain their attention. To further address this problem, program sessions were split from three straight days into six 3-hour workshops spread over a two-week period,⁶ which also gave them a chance to reflect and internalize the skills learned through homework over the weekend in the privacy of their own cells.

Results

The aim of this pilot study was to evaluate the relevance and usefulness of TOC training for inmates in helping them reintegrate into the workforce, which was to be measured quantitatively through job retention upon release, and qualitatively through inmate feedback. Results were compiled by the *Singapore Corporation of Rehabilitative Enterprises* (SCORE), the rehabilitative division of Singapore Prison Services, fully independent of TOC, upon completion of the pilot project.

Quantitative

Success of the project was measured quantitatively by job retention upon release of our sample population. According to data provided by SCORE as illustrated in [Fig. 27-12](#), historically approximately 20 percent of inmates were able to retain a job upon release for three months or more. At the end of the pilot study, job retention over a three-month period rose threefold to 59 percent.

Qualitative

Inmate Feedback

Surveys were administered to inmates on the last day of the training program to appraise their perceptions of the usefulness of the training course, and to evaluate if the tools learned were being utilized upon completion of the course. An evaluation form consisting of both closed (Section I) and open-ended (Section II) questions was prepared for this purpose as seen in [Fig. 27-13](#).

Cluster A Job Fair - TOC Participants Update					
	PROG	DR			Total
Total TOC applicants (Nov)	30	36			66
Did not apply	0	6			6
Successful	30	30			60
Unsuccessful		0			0
Unable to start due to national service or extended imprisonment		1			
Pending					0
Outcome of job TOC applicants	30	29			59
(a) Started work	30	6			36
(b) Found own job		13			13
(c) Yet to take up job		0			0
(d) Declined job offer		5			5
(e) No response yet		5			5
(f) Yet to be released from prison		0			0
TOC tracking update	Overall			%	
Stay on job for 3 months (Nov/Dec/Jan)	1st Month	36		60	
	2nd Month	36		60	
	3rd Month	34		59	

FIGURE 27-12 Results of TOC pilot study.

In addition, the 59 inmates who completed the evaluation gave the course a mean rating of 5.75 from a total of 6 for 13 questions rated on a Likert scale in [Fig. 27-14](#), suggesting that inmates found the training very relevant and useful. Of the different aspects evaluated, inmates were most enthusiastic about the usefulness of the tools and how they helped them distinguish between needs and wants, overcome obstacles to achieve their goals, and understand flawed assumptions.

		<i>Strongly Disagree =</i> 1			<i>Strongly Agree =</i> 6			
		1	2	3	4	5	6	Average
Q1	I still think the course is useful to me.	0	0	0	1	7	50	5.84
Q2	The course helps me understand the differences between needs and wants.	0	0	3	0	4	51	5.88
Q3	I am using the tools and techniques I learned during the course to change the way I think about problems.	0	0	0	1	12	45	5.88
Q4	I am using the tools and techniques I learned during the course to help me see different perspectives.	0	0	0	4	13	41	5.64
Q5	The course helped me understand the importance of questioning my assumptions.	0	0	0	4	14	40	5.64
Q6	I am using the tools and techniques I learned during the course to question my assumptions.	0	0	0	1	15	42	5.7

Q7	The course helped me understand the results of flawed logic.	0	0	0	3	14	41	5.65
Q8	I am using the tools and techniques I learned during the course to correct wrong assumptions and logic.	0	0	0	2	14	42	5.69
Q9	I am using the tools and techniques I learned during the course to help me overcome barriers towards achieving my goals.	0	0	0	1	13	44	5.74
Q10	I am using the tools and techniques I learned during the course to help me achieve ambitious goals.	0	0	0	2	14	42	5.69
Q11	I am more positive about my ability to achieve my goals after attending the course.	0	0	0	2	10	46	5.76
Q12	I have a better understanding of the realities of the Singapore job market.	3	1	3	3	21	46	5.05
Q13	I would recommend other inmates to attend this course.	0	0	1	0	5	52	5.85

74.01

FIGURE 27-13 Inmate evaluation form.

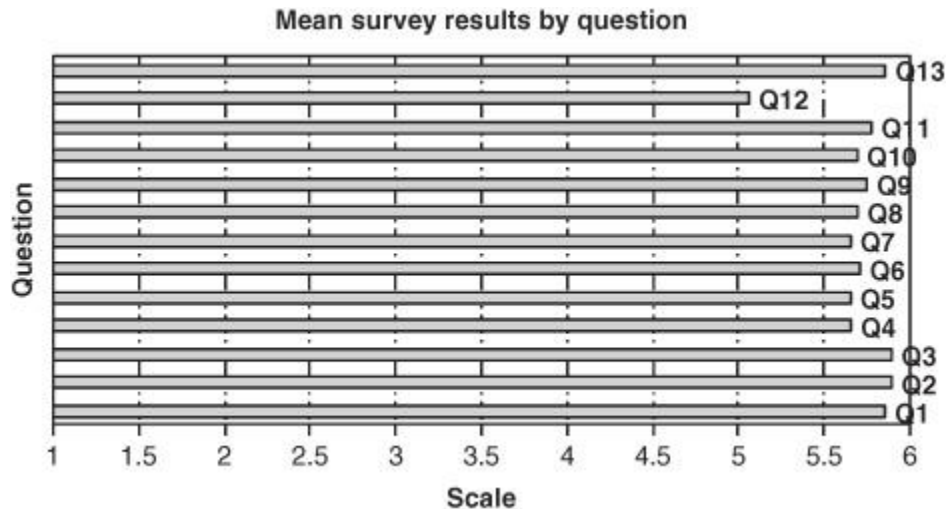


FIGURE 27-14 Inmate Likert scale.

Responses to open-ended questions were congruent with their responses to their ratings in Section I. Inmates expressed that TOC helped them overcome internal conflicts, plan “one-step-at-a-time,” make better decisions, set realistic goals, and focus on the future. Other comments from inmates include:

- This is the best thing that has happened to me in prison. If I attended the course during my first sentence, I would not have come back to prison again.
- Most important it helps us to *think wisely* (which many of us do not usually do) and ask ourselves *why* (which we never thought of asking) and to really identify our basic needs.
- It enables me to discover the missing pieces of the puzzle that I’ve been searching for all these years. At last I got an answer to my agony.
- It will help me think of consequences in doing what I think is safe.
- Creates a clear picture of why I have failed so many times so that I can better understand where my flaws are.

- It really makes me understand myself much better. . . . but the truth sometimes hurts.

Trainer Feedback

“Even though we also experienced deprivation of sorts while we worked in a high security setting, it is great to know that the TOC tools gave them freedom even before they were being released. They were hopeful because their thinking had begun and were able to reason within themselves. Some broke free from mindsets and their behavior changed immediately.

“One found a way to prevent the violence that led him to prison. This discovery alone brought a smile to his face that he could finally control himself. Upon release, another was able to stay in a physically demanding job just by thinking through and deciding what is more important. Another took two jobs so that he could keep himself busy and avoid being bored, which he knew would lead him to trouble. This same person now has an Ambitious Target of wooing his ex-wife and he has hope. These little reality stories affirm that empowering the ‘scum of society’ with tools that work gives the most meaningful satisfaction.”

Government Feedback

“This programme uses simple and logical tools to question the assumptions underlying conflicts in one’s life so as to enable one to make correct decisions in life. This is useful for offenders as it helps them to remove their self-created barriers towards employment.”

—*Speech by Mr. Zainul Abidin Rasheed, Senior Minister of State, Minister of Foreign Affairs at Yellow*

Ribbon Job Fair on November 1, 2006 at Changi Prison.

Judging from the measurement criteria provided, we can conclude that this pilot study was a resounding success at the point of measurement and provided a solid foundation toward behavioral modification for inmates. Both the high job retention rate and the excellent feedback surveys indicate a strong buy-in for TOC, making it a worthwhile cause for further work in prison rehabilitation. Much of the coursework submitted by the inmates showed clear resolution to their long-term chronic conflicts and negative behavior, thus reinforcing the effectiveness of TOC as a reflective tool for inmates to analyze their past actions.

Longer term, using these same measurement criteria, it is impossible to empirically evaluate the effectiveness of the program in bringing about and maintaining the desired change; as change can only occur if ex-inmates continuously practice the skills and use the tools learned, which is impossible to follow up and measure post-release. On a purely subjective basis, however, we have seen enormous transformations in those ex-inmates who chose to stay in regular contact over the last two years and shared their testimonies.

Outside the parameters of this study, a better gauge of long-term success may be to measure the average time between reoffense or recidivism for multiple offenders. While the ambitious target may be to achieve 100 percent rehabilitation, it is neither a possible nor a realistic target due to overwhelming social pressure upon release. A simplistic and more pragmatic challenge given to the inmates during the course is to stop and ask the question "why" in the TOC context before any negative action is taken. Even if it just slows down the time between crimes by one day or one year, it is still an exceedingly worthwhile exercise by helping to restart the logical thinking process.

Follow-on Implementations

Subsequent to the success of the initial pilot project, further trainings have been completed for:

- Other pre-release inmates
- Young offenders (aged 14 to 21 years old)
- Prison officers
- Prison counselors and psychologists
- After-care workers

Future Recommendations

Going forward, the following factors should be taken into consideration.

Delivery. Much of the inmate feedback centered on introducing the course earlier into the incarceration period, which would give more time to absorb and reflect on concepts taught. Due to funding restrictions, the pilot was performed during the final prerelease stage of incarceration, which coincided with a flux of other prerelease programs. Timing was very tight and did not allow for much learning reinforcement and those requiring additional coaching. Earlier introduction during the treatment phase spread over more training sessions would be more beneficial through more practice under supervision.

Follow-up courses. Lack of follow-up is a major problem after behavior training. While marked change was observed during the program, the tendency is to resort to default learned behavior without supervision. For example, in the case of a drug addict, there are many other new problems to deal with once they are sober. This requires readiness to face reality and to take responsibility for their actions.

Follow-up programs available upon release would provide direction and support for ongoing family and work problems outside prison, and act as a valuable network for participants to meet regularly and discuss mutual problems similar to the Alcoholics Anonymous framework.

Family inclusion. After an inmate is released, the family plays a very important part in helping him stay away from crime. However, these familial bonds are stressed greatly when the inmate is incarcerated. Upon release, even though the intention is to help, the result is often misunderstood. A modified version of the inmate training for family members would help to create better awareness and understanding of the inmate's issues and perspectives as well as provide support and reinforce his learning and behavioral changes.

Measurement. The existing measurement criterion—that is, job retention rate—is heavily influenced by many external factors that are outside the control of the study. Introducing the program at an earlier treatment phase of incarceration would allow measurement of behavior within the controlled prison environment that is more accurate.

Summary and Conclusion

The main purpose in writing this chapter is not only to share the wonderful achievements made by our “boys,” but also to encourage you, the reader, to explore the many opportunities waiting to be unlocked by TOC. Looking back, our first challenge was how to implement TOC within a society as efficient as Singapore. The answer is that TOC is so ubiquitous that it can be used to benefit all. We just need to open our minds and look for opportunities outside traditional industries. In this regard, instead of asking “why?”, perhaps we should be questioning our assumptions as to “why not?”

In many ways, designing this course has been a reawakening to the powerful yet elegant simplicity of the core concepts underpinning TOC. If the goal of TOC is to teach the world to think, then we need to dissect what is at the core of TOC and impart that to the masses. Compared to many corporate clients who can demonstrate technically perfect trees, many inmates walked away from the course with just a vague memory of the sequence or terminology of the actual TOC process steps. Having applied TOC to their personal conflicts and daring to honestly question and examine their underlying logical thinking processes, however, the core principles and essence of TOC will remain with them for a lifetime.

Currently we are working on a project with the Singapore Ministry of Community Development, Youth and Sport to work with *People with Disabilities* (PWDs) to achieve their full potential and function independently by overcoming their personal constraints using TOC TP tools. On a completely different level, we recently delivered a professional development TOC TP seminar to CPA Australia accountants. Regardless of whether the audience is an inmate, disabled, or a white-collar professional, it is clear that these same TOC TP tools can equally impact lives at any level. By having the opportunity to stand on the shoulders of giants, it is indeed a privilege for us, as TOC practitioners, to be able to take others less fortunate along for the view.

About the Author

Christina Cheng is the Singapore Director for *TOC for Education* (TOCfE) and runs her own consultancy business. She has spearheaded several new TOCfE initiatives outside of the traditional school framework and has trained diverse adult and youth audiences within government organizations

and the social services sector in Singapore using TOC in behavior applications.

Before her involvement with TOCfE, Christina was a financier and private equity manager with a major European bank covering the North and South East Asian markets. She is an Australian, married with two children, and resides in Singapore.

SECTION VII

TOC in Services

CHAPTER 28

Services Management

CHAPTER 29

Theory of Constraints in Professional, Scientific, and Technical Services

CHAPTER 30

Customer Support Services According to TOC

CHAPTER 31

Viable Vision for Health Care Systems

CHAPTER 32

TOC for Large-Scale Healthcare Systems

As a burgeoning segment of many economies, Services offer incredible opportunities for improvement using TOC concepts. How the tools of TOC work for these environments and how to implement them is discussed for technical, scientific and professional services, customer service support, medical practice, and hospitals. Recall in the last section on Thinking Processes we had a chapter on the use of TOC in education and a chapter on the use of TOC in prisons. The application of TOC in these areas has produced dramatic results, as the chapters in this section will show. These chapters feature the use of Critical Chain, Buffer Management, and the TOC Thinking Processes explained in

earlier sections to great effect in services environments. Services impact on company revenue and profit is challenged by shortened product life cycles, and increasing process and product complexity. TOC methods help ensure that services actions flow in a manner to support escalating demand for services while at the same time retaining financial viability for services within the business.

CHAPTER 28

Services Management

Boaz Ronen and Shimeon Pass

Introduction

Service organizations usually strive to excel in the professional or technical aspects of the services they provide to customers. However, managerial improvements have a huge potential for enhancing shareholders' value of the typical service organization. In this chapter, *value enhancement* will be the main criterion for examining the potential and importance of the traditional versus the more modern managerial concepts and tools. In business organizations, the firm's value is defined as the discounted cash flow (Ronen and Pass, 2008a). In non-profit organizations, the goal is to increase the relevant performance measures versus the organization's goal (Ronen et al., 2006).

In order to improve the performance and value of an organization, we identify its main *value drivers* (Ronen and Pass, 2008a, [Chapter 19](#)). A *value driver* is any important factor that significantly affects the value of the firm. The potential value drivers are identified by a focused review and analysis of the organization. In service organizations, typical value drivers are increasing sales Throughput, increasing *information technology* (IT) Throughput, reducing lead times, and changing measures of performance.

The scope of our discussion on Service Management covers organizations such as:

- Banks (Ronen and Pass, 2007)
- Insurance companies (Eden and Ronen, 2007)
- Cellular phone operators and providers (Ronen and Pass, 2008b)
- Telcos (Ronen and Pass, 2008b)
- Credit card companies (Geri and Ronen, 2005)
- Hospitals and health care service providers (Ronen et al., 2006)
- Law courts
- Professional services: law, accounting, consultation, engineering, design, IT consulting, etc. (Ronen and Pass, 2008a)
- Retail companies and chains (Ronen and Pass, 2008a)

Copyright © 2010 by Boaz Ronen and Shimeon Pass.

- Hospitality industry
- Education (Goldratt and Weiss, 2006)

Challenges in Service Management

In the beginning of human civilization, people were struggling against the hardships of life—most of the population was involved in the production of food, housing, clothing, and defense against enemies. The advancement of civilization relieved the hardships of life, the standard of living has been elevated, and today most people provide services to other people.

Approximately 80 to 95 percent of the global workforce is employed in the service sector. The other 5 to 20 percent

work in either the manufacturing or the agriculture sectors. In contrast to the high importance of the service sector to the global economy, and unlike the production sector, management practices used by the service organizations are not necessarily the state-of-the-art management practices.

From the authors' experience in implementing *Theory of Constraints* (TOC) and other value-focused concepts and tools in dozens of service organizations worldwide, significant changes can be achieved quite easily for the benefit of all the service organization's stakeholders.

This gap in the use of management practices between the production and the service sectors stem from several reasons. We will first focus on factors that make the service sector different.

What Makes Service Management Distinctive?

Service management has several unique characteristics:

- The outcome of a service is not physical in its nature.
- There is a large variance among service organizations (even within an industry) in terms of customers, service types, service providers, and service procedures.
- The goal of service organizations is not always clear, particularly in nonprofit organizations.
- Measurement and control are not trivial.
- In service, the customer is often part of the process.
- Service cannot be made in advance or stored as inventory.
- Entities within the service process are not always visible or physical.
- Bottlenecks within the service processes are, in many cases, hard to detect.

- Many of the service organizations are nonprofit organizations.
- Service organizations are usually labor intensive.
- In many service industries, operations and core processes require high levels of IT capabilities. In these organizations, IT applications development resources are *permanent bottlenecks* (Pass and Ronen, 2003).
- In most service organizations, there exists a high percentage of fixed costs—usually much higher than in a manufacturing firm.

One might get the impression that because of these characteristics, service organizations cannot utilize TOC and other practices developed by the manufacturing sector. The next paragraphs will show that this is not the case. Moreover, due to the existing gap in the implementation, service organizations have a huge potential for value enhancement.

Why the Need for Change?

Most service organizations lag behind the progress that industrial organizations have made in implementing new managerial management methods, such as TOC, Lean/Six Sigma, or *Total Quality Management* (TQM).

Most service organizations have not yet assimilated the understanding that they can leverage excellence in operations to increase shareholders' value. Similarly, the notion of service quality is sometimes misinterpreted. In many cases, service organizations put some effort on Lean/Six Sigma implementations mainly in the operations. Usually, this effort does not improve the organization's value in any significant way.

Proper management of the system's bottlenecks, a change in local performance measures, lead time reduction, decision making, and pricing or costing procedures are major opportunities for the improvement of service organizations.

This chapter aims at presenting the state-of-the-art management concepts and tools, demonstrating their potential for value enhancement for service organizations. This chapter will also suggest proven routes for value driver identification and successful implementation in a variety of service environments.

Second, this chapter surveys the literature on TOC in service organizations. Third, a brief assessment of service management is presented. Fourth, concepts and tools of TOC and Focused Management for service organizations are described. Fifth, an implementation plan for service organizations is presented. Sixth, the remaining chapters of the services section are listed.

Survey of Service Organizations TOC Literature

Literature Mapping and Observations

Relative to the spread of TOC literature in manufacturing, logistics, and project management, little research has been conducted on TOC in services. In addition, there exist only a few papers describing TOC and Focused Management implementations in services.

Unlike manufacturing, project management, or distribution, the service environment has much higher variation. A bank is different from a production organization in its processes, information flow, and core problems. In general, we can classify manufacturing plants into V-, A-, or

T-plants. In service organizations, the variation is much higher.

Some observations have been made from surveying the literature and are detailed in the following section.

From All Service Industries, TOC Is Relatively Most Popular in Healthcare Organizations

Of all service industries, TOC is relatively most popular in health care organizations. The reason might be the fact that hospitals, clinics, and other healthcare organizations are “production lines” dealing with billions of people per year. Some departments are, in fact, job shops. Others are V-, A-, or T-plants, and many can be considered as project-like sites. They have bottlenecks, and their *work-in-process* (WIP) can be easily seen. Measures of performance are operations-like. All the issues in which TOC has proved its ability to improve are in the nature of healthcare organizations. Thus, there exist “full-scale” implementations and methodologies in health care organizations. Ronen et al. (2006) prescribed an end-to-end methodology based on TOC and Focused Management methods that has significantly increased Throughput, reduced lead time, and improved quality in health service organizations with existing resources. Motwani et al. (1996) illustrate how TOC can be applied to service and not-for-profit organizations. Umble and Umble (2006) describe a successful implementation of buffer management in the UK national healthcare system. This research illustrates recent applications in the Accident and Emergency departments and the hospital admission process of three facilities.

Wright and King (2006) describe the problems and the environment of a health service organization in a novel (a *The Goal*-like book) *We All Fall Down: Goldratt's Theory of Constraints for Healthcare Systems*. The issue of

implementing TOC in a hospital has inspired the leading healthcare community, and the book, although not a scientific one, was presented in the prestigious *New England Journal of Medicine* (Pauker, 2006).

Young et al. (2004) describe three established industrial approaches—Lean thinking, TOC, and Six Sigma, and explore how the concepts underlying each of them might relate to health care. Leshno and Ronen (2001) described the complete kit concept as a part of a full Focused Management implementation (constraint management, WIP reduction, performance measures alignment, and strategy) in a private hospital. Ritson and Waterfield (2005) present a case where TOC was implemented in a mental health service.

“One TOC Tool” Implementations and Research

Except for the healthcare industry, where several TOC tools were implemented, in all other service industries we observe “one TOC tool” implementations.

For example, papers were focused on the application of *Throughput Accounting* (TA) or the elimination of prevailing costing practices in the service organization. Roybal et al. (1999) focused on using activity-based costing and the theory of constraints to guide continuous improvement in managed care. Gupta et al. (1997) integrated TOC and *Activity-Based Management* (ABM) in a health care company.

Patwardhan et al. (2006) used the TOC tool of the *thinking processes* (TP) in Evidence-Based Practice Centers.

A Large Part of the Literature Is Examining the Feasibility of TOC Applications in the Service Industry

Goodrich (2008) has explored the potential of using TOC in change management for professional service organizations.

Taylor and Churchwell (2003) have investigated the feasibility of the TP and their potential in a state hospital. Schoemaker and Reid (2005) explored the use of TOC TP and applied it the government sector, at the Albuquerque Public Works Department. Reid and Cormier (2003) applied the TOC TP in the services. Moss (2002) has explored the feasibility of using main TOC tools in service firms.

Limitations of Current Research

TOC research lags behind the research done on other managerial methods. A quick and non-scientific literature survey using Google Scholar (2009) reveals that the citations of TOC-related topics are by far fewer than those on TQM and Lean/JIT. For example, the term “Theory of Constraints” is cited 6680 times, as opposed to 23,700 for “Lean Production” citations, and 281,000 for “Just in Time.” “Drum Buffer Rope” is cited 906 times, while “Kanban” is cited 18,900 times. Goldratt is cited 6300 times while Deming got 142,000 citations.

So, what are the core problems of TOC research?

TOC is a simple and practical tool for better management. Simplicity is seldom a main desire for current business academic research. TOC does not use any complex stochastic and deterministic models. Rather, it uses heuristics that work well in practice.

The academic TOC community is relatively small, as TOC is not yet the mainstream in management. The main performance measures of an academic person are the quality and the amount of his or her research. TOC research does not support the route to tenure track for a young researcher.

Last but not least, the TOC community is a closed community of people getting their knowledge from a limited amount of sources. In the last few years, TOC has

concentrated mainly on the *Viable Vision* (VV) projects in production and logistic organizations, and putting efforts on *Critical Chain Project Management* (CCPM). Thus, the important issues of how to implement TOC in service and nonprofit organizations were lagging behind.

Brief Assessment of Service Management

What to Change?

The value drivers for service organizations that hold a great potential for value improvement are:

- Proper definition of the goal
- Measurement and control
- Constraints management, especially in the IT department
- Emphasis on shortening lead times and improvement in *due-date performance* (DDP)
- Proper decision making especially regarding pricing, costing, and transfer prices
- Proper management of the Sales and Marketing departments

The service industries where the WIP is physical (like retail or health care, since WIP is comprised of customers or patients) were in fact the first to implement some of the concepts and tools discussed in this section. However, the need to change is especially prominent in organizations where the WIP of the service is non-physical (e.g., software code, requests for life insurance policies).

Why Is TOC Not Yet Popular Among Service Organizations' Managers?

TOC is less popular in service organizations than it is in production management. There are several reasons for this gap:

The “production/manufacturing” language— To most service organizations’ managers, some topics seem to be relevant only to the production world. “Batch size,” “load,” “setup,” “Throughput,” “cost per unit,” “complete kit,” “buffer,” etc. seem to them as not applicable to the service environment. As matter of fact, all of these issues are also highly relevant to service organizations.

Lack of immediate quick wins in operations— TOC and Focused Management received their popularity in the manufacturing sector due to the fact that they were able to achieve substantial improvement in operations in a relatively short time. Many of the improvement areas that brought quick wins in terms of value enhancement have a lesser effect or are difficult to achieve in the service environment.

WIP-related problems are more difficult to resolve — In the implementation of TOC in production, WIP is substantially reduced, with profound effects on performance and value. Smart scheduling procedures and *Drum-Buffer-Rope* (DBR) implementations were delivering the “miracles.” WIP in service is also a major problem, yet it is more difficult to resolve. This is especially true in service industries where WIP is non-physical.

No raw materials and finished goods success stories— As service organizations do not have *raw material* (RM) and *finished goods* (FG) in their core processes, the proven TOC methods for these areas do not apply in service organizations.

Bottlenecks are usually not easy to identify—In the service environment, bottlenecks are not visible. This

is especially true for service industries where WIP is virtual.

Lack of *abody of knowledge*(BOK) and experience on how to deal with service organizations—

Production companies are very similar to each other. Practices and procedures were developed along the years to deal with V-, A-, and T-plants. Since service organizations have high variation in processes, structure, and workflow, there are no generic practices for their improvement. Unfortunately, the *TOC International Certification Organization* (TOCICO) BOK in recent years mainly has focused on production issues, concepts, and tools.

VV projects are focused on production, logistics, and manufacturing.

Difficulties in defining the goals of nonprofit service organizations— The lack of a clear definition of the goal in nonprofit service organizations blocks the successful implementation of improvement projects.

For nonprofit organizations, there are difficulties in measuring performance and TOC is perceived as a business-oriented philosophy.

What Do TOC and Focused Management Have to Offer?

Our experience in implementing TOC and Focused Management concepts and techniques show that despite the difficulties listed previously, there exist tools, practices, and methodologies that bring about major improvements for service organizations.

Later in this chapter, we will present concepts and tools for successful implementation in service organizations.

TOC Concepts and Tools for Service Organizations

This section describes a coherent methodology for managing service organizations based on TOC literature and the experience of the authors in implementing TOC in dozens of service organizations of different kinds.

The Seven Focusing Steps of TOC

The seven focusing steps of TOC (Pass and Ronen, 2003) form a very effective framework for managing service organizations. The seven-step framework adds two preliminary steps to the common five-step framework introduced by Goldratt (Goldratt and Cox, 1992). The first step deals with the definition of the goal, while the second step deals with the definition of a corresponding set of performance measures. The addition of these first two steps is highly important for nonprofit service organizations. Thus, the Five Focusing Steps (5FS) framework comprises the following steps:

1. State the goal of the organization.
2. Define global performance measures.
3. Identify the system constraints.
4. Decide how to exploit the system constraints.
5. subordinate everything else to the constraints, and to the above decisions.
6. Elevate the system's constraints.
7. If a constraint has been broken, go back to Step 3.
Warning: Do not let inertia become the system's constraint.

In the first step, the goal of the organization is defined. The goal of profit organizations is to increase shareholders' value. Shareholders' value is the discounted cash flow of the organization. The value-centered definition of the goal is important in service organizations because it focuses everybody on the organization's value. In nonprofit organizations, focusing the whole organization on the goal is even more important because the definition of the goal is usually more complicated and requires incorporation of the resources limitations typical to many nonprofit organizations.

In the second step, a set of performance measures is defined for the organization and its units. Performance measurement is not widespread in service organization but its importance is high—it serves as a compass for management to monitor and control how the organization is functioning and, eventually, the achievement of its goal.

Bottleneck Management

Similar to most organizations, the overall constraint of service organizations is the *market constraint*. Namely, many of them have excess capacity for selling more of their services and the ability of earning more money is governed by the demands of the market. In other words, service organizations are able to cope with a substantial increase in the number of customers and serve them properly. This becomes even more important because most of the costs of a typical service organization are fixed costs.

As noted by Pass and Ronen (2003), most service organizations have two bottlenecks: one in their Sales and Marketing departments and the other in their IT department. These internal bottlenecks will remain bottlenecks even if more resources are added to the respective departments and therefore are referred to as *permanent bottlenecks*. The

IT department is the heart of service organizations such as banks, insurance companies, cellular providers, telecommunication firms, and credit cards companies.

No matter how many more salespersons or marketing employees we add to the Sales and Marketing departments, there will *always* be more potential customers and sales and more marketing initiatives and activities than resources for dealing with them. Similarly, no matter how many development people we add to the IT department, there will *always* be more demand by other departments for development projects and enhancements to existing information systems. This demand for development in the IT department is typically in excess of 300 to 500 percent compared to their development capacity.

The operations, logistics, and customer service departments should not be bottlenecks and these departments should be planned and run with a proper amount of *protective capacity* (Ronen and Pass, 2008a, [Chapter 14](#)). *Protective capacity is a controlled excess capacity aimed at protecting the undisturbed flow of service transactions through the organization.*

A diagram that is useful for illustrating this permanent bottleneck reality is called *Cost-Utilization* (CUT) diagram (Ronen and Spector, 1992). The CUT diagram is a histogram that schematically compares the utilization (load) of each resource of the organization with its cost. Each bar in the histogram represents a single resource or department; the height of the bar corresponds to its load (0 to 100 percent), while its width is proportional to its cost.

Focusing management on the permanent bottlenecks has an immense potential of substantially enhancing the performance and value of the organization. Improved performance of Sales and Marketing will bring more customers, whereas improved performance of the IT development department will allow offering better service to customers.

Exploiting Permanent Bottlenecks

Since the service organization always has permanent bottlenecks at the Marketing and Sales department as well as at the IT department, these bottlenecks should be properly managed in order to secure their best *exploitation*.

Exploitation of the bottleneck has two dimensions:

1. Efficiency—reducing nonproductive times of the bottleneck
2. Effectiveness—directing the bottleneck to process the most valuable services, tasks, and customers.

Increasing Bottleneck Efficiency

Although the value of the service organization is highly dependent on the outcome of its bottlenecks, the percentage of the time that bottlenecks are productive is much lower than 100 percent—usually in the range of 40 to 80 percent (Ronen and Pass, 2008a, [Chapter 17](#)). The nonproductive time is called *garbage time*. *Garbage time of a bottleneck is the time devoted to activities that either nobody should be doing or surely should be done by another (non-bottleneck) resource.* The garbage time is caused by activities such as rework due to an incomplete kit of requirements or instructions, and participation in unnecessary meetings.

Reduction of the garbage time is achieved by a simple procedure: monitoring the wasted times, classifying them according to their causes, using the Pareto analysis to identify the main causes, and implementing remedies that eliminate or greatly reduce the main causes of wasted time (Ronen and Pass, 2008a, [Chapter 5](#)).

The typical result of such a procedure is a 20 to 40 percent increase in the bottleneck's Throughput. Namely, by conducting an easy-to-implement procedure one can get potentially 20 to 40 percent more salespersons or more software developers without any investment in expensive salaries or training.

Increasing Bottleneck Effectiveness

By definition, bottlenecks are resources that are not able to perform all tasks arriving at their desk. Instead of letting chance dictate which tasks will be carried out and which will be abandoned, it is much wiser to choose to accomplish those tasks that will bring most value to the service organization and abandon the least valuable tasks. *The systematic process of picking the most valuable tasks for execution is called strategic gating* (Pass and Ronen, 2003). Strategic gating is a process of prioritization that defines the value of the different tasks, products, services, projects, or customers for the organization and decides by priority which ones will be carried out and which will be dropped (Ronen and Pass, 2008a, [Chapter 5](#)).

Priority of a task/product/service/project/customer is affected by two parameters—on one hand, its value to the organization and on the other hand, the time (effort) spent at the bottleneck processing it. The resulting priority can be decided by calculating the *specific Throughput* of the task/product/service/project/customer or graphically by drawing the *focusing matrix* for the bottleneck. *The specific Throughput of a task is the ratio between the value of the task for the organization and the time this task requires to be processed at the bottleneck. Namely, the specific contribution of a task represents the value that the organization gains per constraint time* (Ronen and Pass, 2008a, [Chapter 5](#)). *The focusing matrix is a chart that maps*

the tasks/products/services/projects/customers in two dimensions according to their relative importance to the value of the organization on one hand and the ease of achievement on the other hand (Ronen and Pass, 2008a, [Chapter 5](#)).

In a large financial institute, the total amount of development tasks requested from the IT department was typically 400 percent higher than the actual development capacity. Traditionally, the decision of which tasks to deliver in a given year was influenced mainly by the organizational power of the requesting unit (“he who shouts louder, wins”). In order to decide rationally on the best portfolio of tasks to be developed during the next year, management adapted and implemented the strategic gating mechanism. A major element of strategic gating is the notion that those tasks that did not have high enough priority should not be put on a “contingency list” but will be put aside in a *firm freeze* status waiting for a subsequent annual strategic gating session.

This strategic gating process obviously ensured that maximum value to the organization was delivered. Moreover, this process increased the effective capacity of the IT department of this firm by 15 percent, enabled it to develop 15 percent more software products, and at the same time enabled it to reduce the damages associated with version content changes.

Subordinating Everybody Else to the Permanent Bottlenecks

The organization as a whole has to be subordinated to its main constraint—the market. This means that in order to achieve high profits one has to offer customers services that deliver as much value as possible.

In order to achieve this subordination to the market, the organization has to undergo a paradigm shift—to accept the

approach that everybody in the organization has to be subordinated internally to the market through Marketing and Sales.

In service organizations, one has to instill double subordination: to the market and to IT development. Subordination to IT development means practically to request only necessary IT applications, to eliminate any “nice to have” features, and to submit the requirements in a complete kit. A complete kit is the set of items needed to complete a given task (e.g., information, drawings, materials, components, documents, tools) (Ronen and Pass, 2008a, [Chapter 12](#)).

Elevating the Permanent Bottlenecks

Permanent bottlenecks obviously can be elevated by hiring more resources. A more challenging mechanism for elevation is the *offload* mechanism. *Offloading bottlenecks is achieved by directing part of the tasks of the bottleneck to other non-bottleneck resources.* Candidates for offload are repetitive tasks or those that do not require the highest professional skills.

Salespersons can be offloaded very effectively by a good back-office. Administrative tasks, meetings coordination, customer retention, etc. can be performed by the back-office, freeing the salesperson to increase Throughput by performing more sales meetings per week. For example, in a medium-sized insurance company, Pareto analysis of a typical day of a salesperson revealed that only 13 percent of the day was used for face-to-face sales meetings with a customers. As a result, they were carrying out only one meeting per day on average. By offloading customer retention activities to the customer support center of the company, salespersons were able to spend twice as much time at sales meetings and conduct two sales meetings a

day. In no time, this trivial change led to a meaningful increase of 20 percent in sales.

Response Time Reduction

Lead time reduction is complementary to the Throughput enhancement by constraint management according to TOC. In order to reduce lead time, it is recommended to implement the *tactical gating* mechanism (Ronen and Pass, 2008a, [Chapter 5](#)). *Tactical gating is the controlled release mechanism for service tasks.* The tactical gating mechanism is based on a “gate keeper” who releases tasks for processing using the following principles:

- DBR scheduling
- Introduction of tasks in complete kit (Ronen and Pass, 2008a, [Chapter 12](#)). For example, in a technical call center in order to serve a customer the service provider needs a complete kit that includes customer name; address; home, office, and mobile phone numbers; name of liaison; details of all equipment on site; nature of failure/complaint; etc.
- Introduction of tasks in small batches (Ronen and Pass, 2008a, [Chapter 11](#))
- Preventing task introduction in an unplanned manner

In order to achieve significant lead time reduction, TOC should be integrated with the tactical gating mechanism. For example, DBR and the complete kit would bring about better results than DBR alone. Addition of the small batch concept (originally suggested by JIT/Lean) and performance measurement would bring about further improvement in performance.

Performance Measures

Goldratt and Cox (1992) suggested three measures of performance for improved management of organizations:

- Throughput (T)
- Operating Expenses (OE)
- Inventory (I)

For service organizations, we suggest to add to this set another three performance measures (Ronen and Pass, 2008a, [Chapter 13](#)):

- Lead time (LT)
- Quality (Q)
- Due-Date Performance (DDP)

Throughput and Operating Expense share the same definition in all types of organizations. In service organizations, inventory is mainly a metric for the amount of WIP in the service process or in a certain department.

Lead time in service organizations should be measured from the customers' standpoint—from the time of the service request by the customer to the moment of service delivery.

Quality is a multifaceted metric. On one hand, the quality perception of the service by the customers is crucial and should be monitored closely by customer satisfaction surveys. On the other hand, the quality of service processes is equally important. The quality of the service process can be monitored by measuring “right first time service,” the amount of “garbage time,” and other industry-specific measures.

Due-date performance measures the adherence of the organization to the Service Level Agreement (SLA) for the

service or the process.

Costing, Pricing, and Decision-Making

Similar to all organizations, service organizations have an improvement potential related to costing, pricing, and decision-making.

The “evils” of traditional cost accounting can be partially resolved by TA. The Focused Management concepts and tools for pricing, costing, and decision-making have succeeded in creating more value in service organizations. For example, the *Global Decision-Making* (GDM) methodology can relieve pricing conflicts, transfer price determination, and make-or-buy decisions as well as investment decisions as shown in Ronen and Pass (2008, [Chapter 16](#)).

Quality Enhancement

Quality is a complicated topic because quality is multifaceted. Service processes are unique in the mere fact that the customer is highly involved in the process.

Some people consider quality to be a cultural issue. In fact, quality is a major business concern having a direct effect on the value of the organization:

- The quality of service processes and the quality of the provided service strongly influence the value perception of customers.
- The quality of service processes has a major effect on the costs of the organization and hence on profits and value. In service organizations, the “garbage time” should be measured by its real economic value, that is, by economic value of the wasted time. The authors’ experience, backed by the relevant literature, shows the

garbage time of IT software developments, for example, is 40 to 70 percent of the labor cost. Methods like “quality costs” usually show a much lower amount of waste and should be scrutinized.

- The business approach to quality encourages the prevention of both poor service quality as well as over-quality for which the customer is neither paying nor valuing (Coman and Ronen, 2009; 2010).

TOC has never developed a coherent methodology for quality improvement in an organization. TOC’s contribution to quality is limited to the issues of where to focus the efforts for quality improvement and the recommendation to eliminate *bad multitasking* (BMT) as a way to improve the quality of execution in project management.

How to Implement the Change?

Value enhancement projects are more complex in service organizations than in manufacturing organizations for reasons mentioned previously. Hence, it is of great importance to have a structured approach to value creation in service organizations. The preferred methodology in this case is the *Value Focused Management* (VFM) approach.

VFM is a practical five-step methodology for implementing shareholders’ value enhancement projects (Ronen and Pass, 2008a, [Chapter 19](#)). *VFM provides a common language across all functional areas; thus, it enables aligning all the organizational decision-making with the goal and creates a clear link between management actions and shareholders’ value.* The stages of VFM are:

1. Define the goal.

2. Determine the performance measures.
3. Identify the value drivers and evaluate their potential impact.
4. Decide how to improve the value drivers.
5. Implement and control.

The buy-in for TOC in service organizations is much more difficult than in manufacturing organizations. Thus, it is important that the management team should be aligned with the goals and methods of the implementation project. Once top management decides to launch the value creation project, thorough education and training seminars should be conducted for top- and mid-management. Value creation teams will then focus on the main value drivers and conduct projects for their improvement (Ronen and Pass, 2008a, [Chapter 23](#)).

The Remaining Chapters in This Section

The remaining chapters in this section are as follows.

Theory of Constraints in Professional, Scientific, and Technical Services by John Ricketts. These types of organizations require the selection of a properly customized portfolio of concepts and tools. Our experience shows that in many cases the team leader or the senior partner is the bottleneck and should be managed as such. In addition, the complete kit concept is crucially important, especially at the beginning of the process.

Customer Support Service According to TOC by Alex Klarman and Richard Klapholz. Customer support units are major elements in organizations in telecommunication, insurance, credit cards, and retail. The unique character of

these units calls for use of the right subset of concepts and tools described in this chapter.

Viable Vision for Health Care Systems by Gary Wadhwa and *TOC for Large-Scale Health Care Systems* by Julie Wright. Physicians usually manage medical organizations, either small or large, because in many countries it is required by law. In this industry, the effect of implementing TOC and other managerial concepts have huge leverage on the performance of these organizations. Not only does it enhance the organization's value but also it substantially improves the quality of the medical service.

References

- Coman, A. and Ronen, B. 2009. "Overdosed management: How excess of excellence begets failure," *Human Systems Management* (forthcoming).
- Coman, A. and Ronen, B. 2010. "Icarus' predicament: Managing the pathologies of overspecification and overdesign," *International Journal of Project Management*, 28 (3), 237-244.
- Geri, N. and Ronen, B. 2005. "Relevance Lost: The Rise and Fall of Activity Based Costing," *Human Systems Management*, 24(2), 133-144.
- Goldratt, E. M. and Cox, J. 1992. *The Goal: A Process of Continuous Improvement*, 2nd ed. Great Barrington, MA: North River Press.
- Goldratt, R. and Weiss, N. 2006. "Significance enhancement of academic achievement through application of the Theory of Constraints." In Ronen B.

(editor). *The Theory of Constraints (TOC): Practice and Research*, IOS Press, Amsterdam, The Netherlands.

Goodrich, D.F. 2008. The relationship of the theory of constraints implementation to change management integration in professional service organizations. Nova Southeastern University, Davie, FL. AAT 3312014.

Google Scholar. 2009.
<http://scholar.google.co.il/> Accessed October 8, 2009.

Gupta, M., Baxendale, S., and McNamara, K. 1997. "Integrating TOC and ABCM in a health care company," *Cost Management* 11(4):23.

Leshno, M. and Ronen, B. 2001. "The complete kit concept—Implementation in the health care system," *Human Systems Management* 20(4):313.

Moss, H. K. 2002. The application of the theory of constraints in service firms. Clemson University, South Carolina. AAT 3057207.

Motwani, J., Klein, D., and Harowitz, R. 1996. "The Theory of Constraints in services: Part 2—Examples from health care," *Managing Service Quality* 6(2):30.

Pass, S. and Ronen, B. 2003. "Managing the market constraint in the hi-tech industry," *International Journal of Production Research* 41(4):713-724.

Patwardhan, M. B., Sarría-Santamera, A., and Matchar, D. B. 2006. "Improving the process of developing technical reports for health care decision-makers: Using the Theory of Constraints in the evidence-based practice centers," *International Journal of Technology Assessment in Health Care* 22(1):26-33.

- Pauker, S. G. 2006. "We all fall down: Goldratt's Theory of Constraints for healthcare systems," *The New England Journal of Medicine* 355(2):218-219.
- Reid, R. A. and Cormier, J. A. 2003. "Applying the TOC TP: A case study in the service sector," *Managing Service Quality* 13(5):349-370.
- Ritson, N. and Waterfield, N. 2005. "Managing change: The Theory of Constraints in the mental health service," *Strategic Change* 14(8):449.
- Ronen B. and Pass S. 2007. "Upgrading the TOC BOK: Focused Methodologies for the Financial Industry," The 5th Worldwide TOCICO Conference, 3-7 November 2007, Las Vegas, NV.
- Ronen B. and Pass, S. 2008a. *Focused Operations Management: Achieving More With Existing Resources*. Hoboken, NJ: John Wiley & Sons.
- Ronen B. and Pass S. 2008b. "Focused Methodologies for the Telco's Industry," The 6th Worldwide TOCICO Conference, 3-4 November 2008, Las Vegas, NV.
- Ronen, B. and Pliskin J.S., with Pass, S. 2006. *Focused Operations Management for Health Service Organizations*, San Francisco, CA: Jossey-Bass (an imprint of J. Wiley & Sons).
- Ronen, B. and Spector, Y. 1992. "Managing system constraints: A cost/utilization approach," *International Journal of Production Research* 24(2):50-53.
- Roybal, H., Baxendale, S.J., and Gupta, M. 1999. "Using activity-based costing and theory of constraints to

guide continuous improvement in managed care,"*Managed Care Quarterly* 7(1):1-10.

Schoemaker, T. E. and Reid, R. A. 2005. "Applying the TOC Thinking Process: A case study in the government sector,"*Human Systems Management* 24(1):21.

Taylor, L. T. III and Churchwell, L. 2003. "Goldratt's thinking process applied to budget constraints of a Texas MHMR facility,"*Journal of Health and Human Services Administration* 26(3/4):416-438.

Umble, M. and Umble, E. J. 2006. "Utilizing buffer management to improve performance in a healthcare environment,"*European Journal of Operational Research* 174(2):1060.

Wright, J. and King R. 2006. *We All Fall Down: Goldratt's Theory of Constraints for Healthcare Systems*, Great Barrington, MA: North River Press.

Young, T., Brailsford, S., Connell, C., Davies, R. et al. 2004. "Using industrial processes prove patient care,"*British Medical Journal* (International edition) 328(7432):162.

About the Authors

Boaz Ronen is a Professor of Technology Management and Value Creation at Tel Aviv University, Faculty of Management. He holds a BSc in Electronics Engineering, and an M.Sc and PhD in Business Administration. Prior to his academic career, he worked for over 10 years in the hi-tech industry. His main areas of interest are focused on firms' value enhancement and TOC.

He has consulted with numerous corporations, healthcare organizations, and government agencies worldwide. During the last 20 years, Prof. Ronen has been leading a team that successfully implemented Focused Management, TOC, and advanced management practices of value creation in dozens of industrial, hi-tech, IT, healthcare, and service organizations.

He has been commended numerous times and received the Rectors' award for outstanding teaching. He was also a visiting professor at the Schools of Business of New York University, Columbia University, Stevens Institute of Technology, several Kellogg programs around the globe, and at SDA-Bocconi (Milan, Italy). Prof. Ronen has published over 100 papers in leading academic and professional journals, and has coauthored four books on Value Creation, TOC, and Focused Management. In 2005, he was the editor of the special issue on TOC published by Human Systems Management. His book on healthcare management was recently published by Jossey-Bass/Wiley. His book, *Focused Management: Doing More with Existing Resources*, was published by John Wiley & Sons in November 2007. Shimeon Pass co-authored both books. His latest book, *Approximately Right, Not Precisely Wrong*, on decision-making, cost accounting, and pricing, was published in 2008.

Shimeon Pass is a noted expert in applying the philosophy and tools of TOC and the Focused Management methodology. He has consulted numerous corporations, organizations, and government agencies worldwide in industrial, service, retail, and nonprofit organizations.

He holds a BSc and an MSc in Chemistry from the Technion, Haifa, Israel and from the Weizman Institute, Israel, and an MBA from Tel Aviv University, Faculty of Management.

Working in the past for IBM in the ERP group, Mr. Pass has also specialized in the implementation of advanced

managerial methods to enterprise information systems.

Mr. Pass is now specializing in applying TOC in the management of R&D organizations and project management.

He has published numerous papers in leading academic and professional journals, and co-authored two books on TOC and Value Creation.

CHAPTER 29

Theory of Constraints in Professional, Scientific, and Technical Services

John Arthur Ricketts

Introduction

Theory of Constraints (TOC) is one of the most widely recognized management innovations of our time. That's quite an achievement, considering TOC creates clear explanations of causes and effects. Although that might sound like TOC is nothing more than common sense, the overwhelming alternative to TOC is conventional wisdom—which is certainly common, but often doesn't make much sense. What sets TOC apart is that it uses knowledge of cause and effect to solve otherwise intractable problems.

For instance, conventional wisdom says the best way to optimize a system is to optimize every element within that system. That's why managers push every worker and every machine to produce as much as possible. Yet, many business and government systems are like chains, and a chain is only as strong as its weakest link: the constraint. Therefore, if the constraint actually limits what a system can produce, conventional wisdom mistakenly calls for lots of process improvement in areas that cannot optimize the enterprise. Indeed, conventional wisdom does not even acknowledge the system constraint, let alone target it for improvement, as TOC does.

TOC is best known in the manufacturing and distribution sectors where it originated, but services are the dominant sectors in mature economies and the fastest growing sectors in emerging economies. Although TOC has been applied in services enterprises, most applications thus far have been limited to services that resemble manufacturing or distribution closely enough that the same applications can be applied. Those applications tend to focus on physical constraints, which are less relevant in most services enterprises, and largely irrelevant in some.

The *Professional, Scientific, and Technical Services* (PSTS) sector is populated mostly by enterprises where physical constraints matter less than intangible constraints. Indeed, the PSTS sector is substantially different, even from other services sectors, for several reasons.

Copyright © 2010 by John Arthur Ricketts.

- Professional, scientific, and technical services are usually customized for individual clients. Repeatability can be elusive when every client wants something different.
- Professionals, scientists, and technicians are highly educated and frequently work in teams. These practitioners have high degrees of autonomy because they are hired by clients for their resourcefulness at solving hard problems.
- Sales are based largely on expertise. Clients expect the experts to have diplomas, licenses, certifications, publications, references, and genuine insights in their fields.
- Delivery depends on intellectual capital, not physical inventory. Know-how is vital in labor-based services. Information technology is vital in asset-based services.

These attributes make PSTS a suitable proving ground for TOC for Services because PSTS is the services sector most different from manufacturing and distribution. Since TOC can work in PSTS, there's a good chance it will work in any services business. This chapter summarizes the adaptation of TOC for PSTS.

The roots of that adaptation extend back to the founding of TOC. *The Goal* (Goldratt and Cox, 1992) is one of the best-selling business books of all time. It tells the story of a beleaguered manager who saves his factory from oblivion—and guides it to prosperity—by applying TOC. It's the seminal work for what's referred to in this chapter as *TOC for Goods* (TOC_G) to distinguish it from *TOC for Services* (TOC_S).

Reaching the Goal (Ricketts, 2008) explains how and why TOC_S differs from TOC_G. It's the foundation for this chapter, but this chapter focuses more on why TOC has taken so long to reach PSTS and what lies ahead in TOC_S.

Background

TOC has been around for decades, so it's reasonable to wonder why it took so long to find an audience in PSTS. The short answer is services are harder to manage than non-services, but unfamiliarity and inertia play big roles as well. So let's start there.

Barriers to Adoption

TOC knowledge is widely accessible in more than 100 books, some of them best-sellers. However, the majority of TOC books are devoted to manufacturing and distribution, while services dominate most economies nowadays. Thus, new readers not only are confronted with unfamiliar TOC

concepts and terminology, but also how those concepts and terms apply to services is an exercise generally left to the reader. And that's not an easy translation, even for TOC experts.

Consequently, TOC is like most management innovations in the sense that it generates more talk than action, but TOC is notable because the action it does generate leads to demonstrable results. Leading manufacturers have applied the TOC application for operations management to varying degrees, and there are pockets among smaller manufacturers, so adoption is far from universal. In services, where the TOC application for project management is the most obvious fit, no more than one in ten project managers use it frequently. The benefits of TOC are extraordinary, however. Improvements of 20 to 50 percent (Mabin and Balderstone, 2000) are well documented, and TOC is thereby a source of strategic advantage.

So why is one of the most-promising management paradigms of our time so hard to adopt? It's a journey that cannot be taken one manager at a time. You've got to take your management team with you. Even if you're a chief executive, you can't do it alone, and you can't just assume that making it a strategic initiative will get the job done. Of course, if you're a manager in a services enterprise, you may have to take your clients and subcontractors with you as well, which makes the TOC journey even more arduous.

It's enough to make any rational manager think twice, then thrice, about taking the leap. Yet some have done so—with success—so it's worth considering the stakes. The Holy Grail of management methods nowadays is process improvement, because competition waits for no one. Therefore, managers are on the lookout for improvement opportunities, to the point that process improvement has become the new business-as-usual.

Typical process improvement methods look for every possible improvement opportunity, on the assumption that

they all add up. However, that's a fallacy because most improvements in a local context create offsetting pain elsewhere that effectively cancels out the benefit. If you look across an enterprise, what you often see is that one manager's pain points are another manager's improvements. Nevertheless, it's not a one-to-one relationship. It's common for one manager's improvement to create pain for tens, hundreds, or even thousands of other managers. That's why process improvements are often thwarted, abandoned, rolled back, or endured grudgingly. The unintended pain of process improvement can be too great for others to bear willingly. When the pain extends to customers, suppliers, and employees, the enterprise can spiral downward, even though the managers pushing local improvements have noble intentions.

Fortunately, process improvement is a domain where TOC really stands out. Rather than casting the widest possible net, TOC concentrates on genuine process improvements by recognizing that an improvement anywhere other than the constraint is a mirage. Making a non-constraint more efficient accomplishes nothing if it further overloads the constraint. And if a local change doesn't move the needle at the enterprise level, it's not really an improvement.

Picture a dozen people all struggling at once to push an enormous crate, with no clear sense of direction or cooperation. Now picture three people easily pulling that crate in unison in just one direction. That's what TOC does.

This can make TOC sound too good to be true. After all, if it really worked that well, wouldn't everyone be doing it? Well, no. What the crate analogy left out is everything it takes to get a team pulling in unison. And the obstacles are formidable.

First, there's the "push, push, push" syndrome. That's the longstanding management mindset that the way we've always done things around here is the way it has to be. Push suppliers. Push schedules. Push workers. Push late jobs.

Push shipments. Push salespeople. Push customers to buy more. In an environment like that, getting managers to adopt a system where things are pulled along naturally sounds as far-fetched as a workable time machine. Besides, they ask, what's a manager to do if there's nothing to push?

Second, there's the "summer love" syndrome. Every management innovation wins some avid converts, but infatuation often fades with the next management fad. The best TOC adoption programs skip the infatuation and go straight to implementations with staying power. To do that, however, you have to know where the real constraint is. And that's harder than it sounds, as we shall see.

Finally, there's the "shoemaker's children" syndrome. This one is particularly acute in PSTS, where every partner, principal, professional, scientist, and technician is an expert in something. If you're a manager in a manufacturing or distribution enterprise seeking to adopt TOC, you're likely to engage an outside TOC expert because their credentials and reputation earn respect among your peers. However, if you're a manager in a PSTS enterprise, those TOC experts can be right down the hall. Not only are they busy doing billable work for your firm's clients, they don't automatically have extra credibility among your peers, who are experts in their own right—just not experts in TOC. Hence, the shoemaker's children syndrome exists when a PSTS enterprise is more successful at helping clients adopt TOC than it is in embracing TOC itself.

Challenges in the PSTS Sector

Some challenges that are endemic to manufacturing and distribution do not carry the same weight in PSTS. For instance, TOC_G strives to minimize inventory because it's an expensive investment that limits flexibility and all-too-often becomes obsolete before it's sold. In PSTS, however, there

are virtually no inventories. Services are consumed as delivered, so there's no way to produce them in advance.

In that context, any solution that minimizes physical inventory is a solution in search of a problem. Nevertheless, as will be seen later, the principles underlying TOC apply to services as well as goods-based businesses and, with some adaptations, TOC_S can address several challenges facing the PSTS sector.

Some challenges facing PSTS are the same as those facing enterprises in other service sectors:

- New entrants have radically different business models.
- Work seeks the lowest level worldwide via outsourcing and offshoring.
- Legislation, regulation, and intellectual property rights can work for you or against you.
- New technology levels the playing field, but old technology is hard to replace.

There are, however, challenges afoot that hit the PSTS sector especially hard:

- Knowledge is expanding, which makes expertise harder to attain.
- The half-life of information is getting shorter, which makes expertise harder to sustain.
- Clients want results, not just advice.
- Demand is inherently unpredictable, so clients want to shift that burden.
- Clients want their projects completed better, faster, cheaper.
- Clients want their processes to accommodate unpredictable swings in demand easily.
- Competitors are observant, so competitive advantages tend to be fleeting.

Fortunately, TOC can address several of these challenges.

What TOC Has to Offer

Challenges facing PSTS are formidable enough to motivate some managers to seek alternatives to conventional wisdom. Fortunately, TOC has much to offer PSTS.

First, TOC establishes flexibility instead of pushing for predictability. That is, rather than striving for more accurate forecasts over longer horizons, TOC manages buffers that anticipate predictable changes in demand or supply. When something unpredictable happens, the enterprise is not locked into lengthy commitments. When you are nimble, rogue waves matter less.

Second, TOC speeds up projects and processes. When correctly harnessed, speed not only makes an enterprise nimble, it pleases clients because they can get their services on demand. Delivering services on demand, rather than as capacity is available, creates competitive advantage that's hard for competitors to match. When you are speedy, everyone else has to play catch-up.

Third, TOC focuses management attention on the constraint. Literally dozens of other concerns can fade into the background when the constraint becomes the center of attention. Moreover, the constraint then becomes a leverage point because relatively modest changes there can generate sizable benefits elsewhere—both for the service provider and its clients. When you manage constraints, noise fades away.

Finally, TOC rearranges management priorities. The top priority for most managers is cost control, but TOC shows how this emphasis is misplaced when it makes growth difficult. In contrast, when managers adopt TOC, their top

priority switches to maximizing cash from sales minus truly variable cost, which is called Throughput. When you maximize Throughput, growth comes naturally.

Every TOC implementation has to answer these fundamental questions: (1) What to change?, (2) What to change to?, and (3) How to cause the change? Answers to these questions for TOC in PSTS are provided next.

What to Change

Pain points would seem to be an obvious way to decide what to change. If you ask managers about their pain points, they can easily reel off long lists. This is in fact the way many process improvement programs actually start. Unfortunately, that sets those programs off in the wrong direction because pain points are symptoms, not causes. Just as treating the symptoms of a disease provides temporary relief rather than a cure, treating pain points provides temporary relief while allowing the core problem to fester.

When applying TOC, *undesirable effects* (UDEs) are the starting point for figuring out what to change. For example, shipping orders late is an UDE of pushing too many jobs into and through a factory. A corresponding UDE in PSTS is finishing client projects late by starting more projects than the service provider can handle at once. Arbitrarily starting fewer jobs or projects isn't the answer, however, because the number of jobs or projects is a symptom, not a cause. Unless you know which jobs or projects to start—and how to manage their constraints—you haven't really solved the problem.

Contrary to conventional wisdom, behind even the most complicated web of symptoms there is usually just a single core problem. It accounts for the multitude of pain points. If there are too many jobs in the factory, the core problem

could be the factory's constraint isn't being managed. If there are too many services projects in progress, the core problem could be the service provider's constraint isn't being managed. The pain points associated with rework, overtime, missed shipments or milestones, employee morale, and customer dissatisfaction can be traced back to this core problem. In addition, mandates to eliminate rework, cut overtime, ship on time, meet milestones, reassure employees, and satisfy customers invoke considerable effort to treat symptoms, not the core problem.

Once the core problem is identified, however, it's usually the result of a conflict. For instance, if senior management complains that utilization is too low, more jobs get pushed into the factory or more services projects get launched. However, the crush of new work and the confusion sown by expediting slow down work on previous commitments, which further depresses utilization. By now, the conflict is in full swing.

The question then should be how to stop the cycle. Conventional wisdom says the way to resolve conflict is compromise. For example, picking an optimal utilization target and setting an optimal production schedule seems like a sensible solution. Unfortunately, universally high utilization and high overall productivity are an inherent conflict, and no amount of compromise will make it go away. Indeed, what often happens is senior management scrutinizes utilization until production becomes unacceptable. Then scrutiny shifts to on-time delivery until utilization becomes unacceptable. Then the cycle repeats. TOC practitioners know, however, that whenever they see an enterprise oscillating this way, its managers are probably compromising on a conflict. Oscillation takes many forms: centralize versus decentralize, hire versus fire, acquire versus divest, and build versus buy, to name a few.

In contrast to conventional wisdom, TOC teaches the way to resolve conflict is to eliminate the conflict itself. For

example, the quest for universally high utilization is rooted in the belief that every resource that isn't fully utilized represents a lost opportunity for production. However, if the non-constraints produce more than the constraint can, work just piles up ahead of the constraint even though non-constraints downstream from the constraint are sometimes starved for work. And if work is released into production just to keep workers or machines busy, it eventually leads to excess inventory. The services equivalent occurs when people bill projects for tasks that could be done better another way, or that don't necessarily need to be done at all to complete the project successfully, but that do contribute to resource utilization.

TOC resolves this conflict by maximizing utilization of the constraint, while minimizing utilization of everything else that isn't required to keep the constraint busy. In other words, the goal is not utilization; the goal is Throughput via saleable products and billable services. Once utilization targets are recognized as the cause of UDEs, no compromise is required to eliminate the conflict—just measure utilization of the constraint, and nothing else.

Not only can service providers use this technique to improve their own enterprise, they can use it to help clients improve theirs. Indeed, the TOC approach to marketing and sales depends on this specific capability. Of course, the TOC approach can be applied to more than just marketing and sales of services. What to change throughout PSTS is covered next.

Expertise and Assets

Every enterprise in the PSTS sector depends on expertise. It's how sales are made and reputations maintained. Missteps here have condemned to oblivion some trusted

professional service firms, cutting-edge research organizations, and high-flying technology start-ups.

If a PSTS enterprise is labor-based, having the right professionals, scientists, or technicians is a critical success factor. Each professional practice, research lab, and technology group needs to have the right skills in the right amount in the right place at the right time.

Conventional approaches include hire-to-plan, which requires a forecast, and hire-to-deal, which requires clients who are patient enough to wait, if necessary. Of course, forecasts are notoriously inaccurate, and clients have less and less patience. Consequently, oscillating between too few and too many resources is a common conflict in PSTS enterprises.

If a PSTS enterprise is asset-based, expertise still plays a vital role. However, the experts put more of their effort into assets that clients value and less effort into serving clients directly. Those assets may be physical capital, such as architectural models, research laboratories, or data centers. Alternatively, the assets may be intellectual capital, such as legal databases, engineering designs, computer patents, or consulting methodologies.

To the degree that assets serve more clients than the experts could without assets, the enterprise gets leverage from its investment in assets. Thus, it might seem that assets lessen the need for experts, but the opposite can be true because a shortage compromises service to multiple clients. For example, a service outage of just a few minutes can provoke howls from all the clients who have come to rely on the service provider's assets.

Service Delivery

Every enterprise in the PSTS sector generates Throughput via projects or processes. Although these terms are

sometimes used interchangeably, making a distinction is useful when applying TOC.

- A project is a set of finite-duration tasks that must be performed in a specified sequence to produce the desired result within a prescribed time and budget, such as designing, building, and implementing an information system. Every project is therefore unique, even if based on a standard methodology with known deliverables.
- A process is a set of activities performed continuously or on a frequently recurring schedule, such as doing legal research, repairing equipment, and processing purchase orders. Every process is therefore highly repeatable, and the output of processes is typically measured in terms of service levels, such as the percentage of service requests completed within a specified period.

If a PSTS enterprise is project-based, it has to execute individual projects, of course. However, it also has to manage a portfolio of projects for multiple clients. Moreover, those projects compete for resources, so project management and resource management are complementary. TOC has traditionally treated resources as relatively fixed, and has managed projects according to the prevailing resource constraint. This approach can be quite acceptable in an enterprise that does internal projects as an adjunct to its main business, such as when it performs engineering projects in support of its manufacturing business.

Although PSTS enterprises also do internal projects, such as those to build their assets, they often do more external projects as their main business. And letting a resource constraint dictate what the service provider can produce may or may not be consistent with its strategy. Indeed, when a service provider adopts a strategy to deliver

services on demand, resources should not be its constraint. Thus, to make TOC workable on such services projects, it cannot be based on the presumption that resources are relatively fixed.

If a PSTS enterprise is process-based, it likewise has to execute individual processes, as well as manage a portfolio of processes for multiple clients. Moreover, those processes compete for resources, not just with other processes, but also with projects. For instance, if the service is employee benefits processing for multiple clients and the service provider is simultaneously building an asset to automate employee benefits processing, then the benefits experts are likely pulled in several directions at once.

Clients engage service providers to perform processes on their behalf for various reasons. Expertise is an obvious one. So is reduced cost from economies of scale. Perhaps less obvious is the expectation that the service provider has global reach, can handle higher processing volumes, or will be able to react to a wider range of demands. The latter point is notable because it requires the service provider to be nimble. The ability to dial processing capacity up and down with demand differentiates services on demand from services as available. Capacity management requires measurements to drive it.

Measurement

Every enterprise in the PSTS sector requires measurement. Of course, the finance and accounting functions are major sources of measurements.

The prevailing measurement method in PSTS, cost accounting, is the same method used in the vast majority of enterprises, regardless of whether they produce goods or deliver services. Despite its widespread use, however, cost accounting is controversial. Many accountants are well

aware of its shortcomings, but they are trapped in a professional conflict that obligates them to use it anyway.

When direct labor costs dominated product costs, allocating overhead was straightforward. However, now that direct labor no longer dominates product costs, allocation creates distortions that mask the true profit contribution of each product. Some products may appear profitable, when actually they are not. Consequently, manufacturers relying on cost accounting make product mix decisions that are far from optimal.

The same dilemma afflicts service providers who rely on cost accounting. Even in labor-based services, cost allocation masks the true profit contribution of service offerings. Some may appear profitable, when they are not. Consequently, service providers relying on cost accounting make service mix decisions that are far from optimal. Moreover, service providers who bid on jobs with cost-plus pricing are more prone to over- or under-price their bids relative to what the work is actually worth to clients.

Another insidious effect of cost accounting is *Cost-World Thinking*, which is the TOC name for making cost reduction the top management priority. Relentlessly driving costs down can have the unintended consequence of driving down revenue, customer satisfaction, and employee morale as well. This is just as true in PSTS as in manufacturing.

Marketing and Sales

Every enterprise in the PSTS sector has to pursue marketing and sales, even if its practitioners are on retainer. Moreover, marketing and sales depends on expertise and intellectual capital that clients value. Here are typical marketing pitches for PSTS.

- We should do it for you because you don't have the necessary expertise in-house (for example, independent auditing, architecture, or intellectual property law).
- We can do it for you because it's not your core competency and we can do it better, faster, cheaper (for example, technical support, procurement, or market research).
- We can do it with you because you have insufficient capacity, need to share risk, require physical facilities, or lack specific skills (for example, joint scientific research).
- We can help you do it yourself by providing assets (for example, information technology, knowledge bases, or patents).

Although it might seem that these marketing pitches, as well as the services they encompass, have little in common, when enterprises in the PSTS sector market such services, they almost always start with cost-plus pricing. That is, they base their bids on standard billing rates, which in turn are based on standard costs plus a standard margin. This, however, assumes that there is only one fair price that all clients ought to be willing to pay.

Of course, standard rates have nothing to do with the business value that clients perceive in a service offer. Two clients receiving identical services may derive substantially different business value because their needs are different. Consequently, negotiations that take place during the sales cycle of large contracts move the provider and the client toward a mutually agreeable price for a given scope of work. How far the provider will negotiate is nevertheless strongly influenced by the margin between standard cost and bid price, which keeps the provider anchored on cost rather than value. Consequently, providers often have no other

way to decide whether they are under- or over-pricing their services.

Something similar happens on smaller contracts, which can include high volumes of services too small to negotiate separately. In that case, the service provider may have discounts based on volume or customer loyalty—and premiums based on local market conditions. Nevertheless, standard rates and margin analysis still lie behind the discounts and premiums, even when there is no overt price negotiation.

The upshot of this is to the degree that standard cost is fallible, the resulting standard rates and gross margin do not maximize Throughput. Furthermore, the nature of the services themselves affects what clients will buy. When every service provider is proposing fundamentally the same services, marketing and sales gravitate to price as a differentiator. This, of course, opens the door to new competitors with different business models that not only change pricing, but also what value clients get for their money.

Strategy

Every enterprise in the PSTS sector tends to have the same fundamental strategy as its competitors. That may sound like a bold statement, but consider this: A typical PSTS strategy says, “the enterprise will provide a given set of services in its fields of expertise to particular types of clients for standard charges—or a negotiated price (within limits based on cost).” It doesn’t matter whether the field is a profession, science, or technology—the strategy is the same.

From the service provider’s perspective, expertise is the primary differentiator, and the tasks are to maintain the firm’s reputation while managing cost, protecting gross

margin, and winning new contracts. From a client's perspective, however, price is the primary differentiator because expertise is essentially unmeasurable. Clients thus ask themselves, "Am I willing to pay a higher price when I cannot objectively evaluate expertise, or can I get reasonably comparable service elsewhere for a lower price?"

This disparity in perspectives opens the door to new entrants who are able to compete only on price. Then the entrenched firms seem to have few choices. They can begin serving different clients with the same services, or existing clients with new services, either of which can shift the battleground to more favorable terrain. Alternatively, they can stand their ground on reputation and hope that their clients are sufficiently risk-averse to low-cost competitors, which means the firm's rainmakers have to cement the firm's relationship with its client base. On the other hand, those entrenched firms can join the price war sparked by the new entrants and watch the market race to the bottom.

There is, however, another possibility: Alter strategy to pursue client value rather than price. The shift from labor-based to asset-based services is one way to do this. It's harder for new entrants to compete on price if they have to build assets comparable to ones the entrenched firms already have. In addition, clients may enjoy the benefits of higher functionality and reliability with asset-based services.

The prevailing PSTS strategy outlined above thus may have made sense when professions were sparsely populated, science was in its heyday, and technology was a novelty. However, with professionals, scientists, and technologists facing competition as never before, an undifferentiated strategy is a huge exposure.

The question then becomes, "Does TOC_S enable service providers to change the game in any way besides shifting

from labor-based to asset-based services?” The answer, as we shall see, is yes.

What to Change to

Let’s get the obvious question out of the way first: Why can’t you just apply traditional TOC to services? The good news is you can, if the services are repeatable enough. For instance, some technical services consist of authorization, delivery or drop-off or dispatch, diagnosis, repair or replacement, shipment or pick-up, and billing. Somewhere among those activities is the constraint, and it can be managed with virtually the same TOC methods used to manage a factory, even if the service provider maintains no parts inventory.

When the services in question do have physical inventory, however, the fit with TOC is even better. For example, food services have to manage not only raw stores, but also *work-in-process* (WIP) in the kitchen, and finished goods on the warming table, in the display case, or on the shelf. Moreover, in some services, things that wouldn’t ordinarily be considered inventory can be treated that way for management purposes. For instance, some health services view hospital beds or operating rooms as finite yet perishable inventory, and manage their processes with TOC. Other health services view each patient’s treatment as a project to be completed within a specified duration (Umble and Umble, 2006).

So if traditional TOC works on some services, why not in PSTS? With a few exceptions, such as the repair service described previously, the services provided by PSTS are not sufficiently repeatable, and inventory is typically a minor consideration. When clients engage lawyers, they want them to win their case. When clients engage scientists, they want them to study their problem. When clients engage

technicians, they want them to fix their technology. Case law, published research, and technical manuals are useful references, but they are not inventory for purposes of applying TOC to PSTS.

Furthermore, services in PSTS are typically customized for individual clients. Even when the service provider has a standard methodology, the services actually delivered have to be tailored to unique customer requirements. For example, when implementing a standard enterprise software package, it has to be configured for the client's information technology environment (servers, storage, communications, firewalls, authentication, etc.), it has to be integrated with the client's other software applications, it has to be loaded with appropriate data files, it has to be tested, and it has to be made usable via demonstrations and training. Although the software itself may be standard, hardly any of the implementation service is actually transferrable between clients.

Consequently, TOC_S is harder than TOC_G for several reasons.

- It can be hard to find the services constraint when there are no piles of inventory to signal where the constraint might be. When services are delivered at client sites or from multiple service centers, you can't just walk around and find the constraint.
- Once the constraint is found, it can be hard to keep the constraint from floating because demand for resources is driven by client engagements. One month the service provider may be short on auditors, the next month short on tax specialists, and the following month short on management accountants.
- Clients are often coproducers of services. Outside legal counsel typically works with inside legal staff. Outside consultants work with company management. Outside

technical specialists work with inside technical staff. Thus, the constraint on services isn't always within the service provider; it can just as easily be within the client's organization.

- There are many sources of variability in services that cannot be buffered with inventory. If the service provider doesn't have sufficient capacity to deliver services on demand, and its clients are unwilling to accept services as available, those clients may choose to do without the service, find another service provider, or do it themselves (Ricketts, 2008, [Chapter 4](#)).
- Service providers may be bound by *Service Level Agreements* (SLAs) that impose penalties on the provider for noncompliance, and may offer bonuses for extraordinary performance. However, if demand for service is outside the service provider's control, such as when the client's customers call contact centers operated by the service provider on the client's behalf, the provider cannot unilaterally deny or delay service without missing the SLA, as traditional TOC applications might dictate.

Despite these difficulties, every one of the applications from TOC_G has been adapted for TOC_S , as will be seen next. Hence, for PSTS the short answer to the question, "What to change to?" is TOC_S .

Replenishment for Services

Replenishment for Goods (R_G) is the traditional TOC application for distribution. Briefly, R_G establishes inventory buffers that cover total consumption of inventory during the time needed to resupply, taking variability into account. Those inventory buffers are located at a central warehouse

rather than retail locations because aggregated demand varies less.

Thus, if it typically takes three to five days to get more of a particular inventory item, and a distributor typically ships 25 units per day, that distributor might size its inventory buffer at 100 units—or 25 units per day times four days. Ideally, this buffer prevents the distributor from running out of that item because the buffer is replenished based on actual consumption. On those rare occasions when the buffer nears depletion, the distributor expedites the active order with its supplier.

Aligning buffer size with consumption and resupply time thus optimizes inventory, so whenever those parameters change, the buffer is resized accordingly. This is a radical departure from conventional wisdom, which says inventory levels ought to be dictated by demand forecasts and infrequent shipments of large, economic order quantities.

Replenishment for Services (R_S) is the TOC application for resource management in services. In general terms, resources are anything a service provider needs to deliver a service, but in labor-based services the term “resources” is virtually synonymous with “people.”

In a services context, resources are not consumed in the same sense as inventory is consumed by distribution. Once shipped, inventory typically does not come back. Once assigned, resources naturally come back for more work. Therefore, R_G is based on total consumption, while R_S is based on net consumption, which is the difference between resources going out on assignment minus those coming back. Net consumption for a given period can therefore be positive, negative, or zero.

Briefly, R_S establishes resource buffers that cover net consumption of resources during the time needed to change supply, considering variability. In addition, those resource buffers are located in skill groups that serve the enterprise

rather than individual projects, because aggregated demand varies less.

Thus, if it typically takes 60 to 90 days to complete the hiring process and get a new employee on board, and the service provider typically needs one additional employee per month in a particular skill group, that service provider might size its resource buffer at two or three resources. Whether it would be two or three depends on whether that particular skill group is the constraint.

No forecasts are needed for R_S , which is a radical departure from conventional wisdom. Thus, R_S is an alternative to hire-to-plan, which begins hiring independent of whatever deals are in the sales pipeline. Likewise, R_S is an alternative to hire-to-deal, which doesn't even start the hiring process until a new engagement is imminent.

Hence, R_G and R_S are based on the same principles, but they operate in different contexts. Furthermore, within services enterprises, R_S supports both projects and processes, which are two different ways to deliver services. They are discussed next.

Critical Chain for Services

Critical Chain for Goods (CC_G) is the traditional TOC application for project management. It's a radical alternative to *Critical Path Method* (CPM) and *Project Evaluation and Review Technique* (PERT), the dominant project management methods.

Briefly, CC_G changes how projects are estimated, planned, performed, and tracked. Those changes lead to better on-time completion as well as shorter projects, which defies the conventional wisdom that says there is an immutable trade-off between those outcomes.

- CC_G employs task estimates at the 50 percent confidence level rather than at the 80 percent level because task-level contingency does not really protect on-time task completion, let alone on-time project completion. CC_G instead starts with non-buffered task estimates, then consolidates contingency into a project time buffer, which protects the project as a whole.
- CC_G eliminates resource contention because overloaded resources cannot complete tasks on time. CC_G instead adds resources or shifts selected tasks earlier in the schedule, which is called resource leveling. The longest path through the project plan after resource leveling is known as the critical chain. Even for projects with equivalent deliverables and scope, the critical chain and critical path are always different because the task durations are different. However, the set of tasks comprising the critical chain and critical path can be different, as well.
- CC_G applies different work rules because projects are best performed like a relay race. That is, each task starts as soon as its predecessors are done, even if that means an early start. Early task starts thus compensate for late finishes elsewhere in the project, which is not something that conventional project management does. Indeed, conventional projects are often late precisely because late task completions are cumulative.
- CC_G tracks projects differently because only a subset of tasks actually determines whether the project as a whole will be completed on time. CC_G measures progress by how much of the project buffer has been depleted by late task completions. If buffer depletion is in proportion to the amount of the project completed (or less), the project as a whole is on track for on-time completion. This contrasts with conventional project

management, which credits every task completion, regardless of whether those tasks actually determine whether the project will be late.

CC_G was invented for engineering projects in a manufacturing context, but it works just fine on individual services projects, too. However, the traditional approach to *Critical Chain Multiple Projects* (CCMP) does not work as well in PSTS enterprises.

CCMP assumes that resources are essentially fixed and that multiple projects must be staggered based on their use of the constrained resource. For instance, if the constrained resource is an aircraft maintenance hangar that holds just one plane at a time, multiple aircraft maintenance projects are constrained by availability of the hangar. The resulting multi-project schedule forms a stair-step pattern based on when each project has an aircraft in the hangar. The same scheduling logic holds, however, if the constrained resource is a person or persons with a particular skill in short supply.

Thus, traditional multi-project critical chain presumes there is an internal constraint. As a result, clients must be willing to accept services as available. However, services clients nowadays are less willing to wait for service. They want services on demand. In addition, some service providers, particularly in the PSTS sector, are anxious to accommodate clients by acquiring additional resources as needed. This changes the multi-project problem to one where there is an external constraint: Either clients will not buy all the services the provider has to offer or the job market cannot supply all the skilled resources the provider needs to meet client demand.

TOC for Services solves the latter problem by combining replenishment with critical chain. That is, *Critical Chain for Services* (CC_S) uses R_S to provide resources to multiple

projects being managed with critical chain. For example, resources on the bench waiting for project assignments—the resource buffer—should be sufficient to meet most demand for resources from multiple projects, even when that demand is unpredictable. Whenever the resource buffer drops below the target size, R_S automatically replenishes it because the resource buffer is there to protect on-time project delivery and the revenue it produces.

By mitigating, if not breaking, the resource constraint, multiple services projects can be scheduled concurrently with CC_S in order to meet the needs of various clients. A stair-step pattern between the projects is therefore not required.

Drum-Buffer-Rope for Services

Drum-Buffer-Rope for Goods (DBR_G) is the traditional TOC application for operations management. Briefly, DBR_G wrings maximum productivity out of manufacturing operations with an internal constraint by ensuring that the constraint, and only the constraint, sets the pace. Indeed, the “drum” in DBR_G refers specifically to this pace setting.

The “buffer” in DBR_G refers to WIP deliberately queued ahead of the constraint. This buffer ensures that the constraint has work to do even when there are disruptions upstream.

As noted earlier, keeping the constraint supplied with work is vital because utilization of the constraint governs what the factory produces overall. Non-constraints therefore must do whatever is required to keep the constraint fully utilized—and no more. That means that non-constraints upstream from the constraint typically have less than full utilization so that they won't overwhelm the constraint with excess WIP. Likewise, non-constraints downstream from the

constraint typically have less than full utilization because they can only do as much work as is passed to them through the constraint.

Late delivery of raw materials, equipment breakdowns, worker absences, unexpected scrap rates, change orders—even the weather—may disrupt the production schedule. Therefore, upstream non-constraints sometimes have to sprint in order to keep the constraint busy when holes appear in the buffer. Likewise, downstream non-constraints sometimes have to sprint in order to complete late jobs on time. Nevertheless, contrary to conventional wisdom, it's normal for non-constraints to be idle occasionally. Indeed, it's necessary for them to be idle at times.

Rather than pushing work into the factory for utilization, DBR_G starts jobs at just the right time and relies on due dates to pull those jobs through the factory in the right order. Thus, the “rope” in DBR_G refers to the information systems used to start jobs at the right time and subsequently ensure that the constraint is working on the right jobs based on current due dates.

Having previously established that PSTS are highly customized and rely little on inventory, how DBR might apply can be a mystery. Nevertheless, it's a mystery that is easily solved.

- When PSTS services are highly customized, the customized processes may nonetheless be highly repeatable. That is, when a service provider uses a shared service center to perform processes for multiple clients, each client's customized process may be performed millions of times. Think of paychecks. Due to organizational structures and compensation plans, no clients' payroll processes are identical, but the same process is performed for every client's employees every pay period.

- Services WIP is often intangible, but it does exist, on paper or in computers. Although such WIP is not exactly inventory, it may nevertheless be managed with similar methods. For instance, service center managers can monitor work queues and completion of service requests. Likewise, research laboratory managers can monitor experiments and completion of milestones.

Briefly, *Drum-Buffer-Rope for Services* (DBR_S) wrings maximum productivity out of service processes with an internal constraint by ensuring that the constraint, and only the constraint, sets the pace. The work buffer ahead of the constraint ensures that the constraint has work to do.

Based on the description of DBR_S thus far, it may seem that DBR_G and DBR_S are indistinguishable, but that is not the case. There is one profound difference: In DBR_G, the manufacturer modulates the work released into the factory in order to keep the constraint busy, while in DBR_S, the service provider cannot control the inputs to the service process. Service requests come from the clients' customers, employees, suppliers, shareholders, or any other group that the client deems eligible for service. The arrival of service requests is not something the service provider can predict with accuracy, let alone control.

If the service provider cannot control inputs to the process, yet is bound by an SLA to deliver service within specified parameters, the process itself cannot have fixed capacity. Hence, the buffer and rope work differently in DBR_S. When the buffer grows beyond its upper threshold, the rope triggers an increase in capacity that eventually brings the buffer level back down into the normal range. When the buffer shrinks beneath its lower threshold, the rope triggers a decrease in capacity that eventually brings the buffer back up into the normal range.

Thus, DBR_G does buffer management of operations with fixed capacity, while DBR_S does capacity management of processes with variable capacity. DBR_G and DBR_S are based on the same principles, but they work differently and are used in different contexts.

Throughput Accounting for Services

Throughput Accounting for Goods (TA_G) is the traditional TOC application for measurement. It's an alternative to cost accounting, the dominant measurement method.

Briefly, TA_G changes financial measures—and therefore other measures derived from them—plus it changes management priorities. One of the financial measures, Throughput, was mentioned earlier. The other financial measures are Investment and Operating Expense (Corbett, 1998).

- *Throughput* (T) is cash from sales minus truly variable cost. Thus, it's revenue minus cost for materials and parts from which each item is built.
- *Investment* (I) is all money invested in things for sale. Plants and inventory are included.
- *Operating Expense* (OE) is all money spent turning I into T. Direct labor, rent, and selling, general and administrative (SG&A) costs are included.

There is no product cost construct in TA. OE is simply summed. It's not allocated to products. This avoids the distortions that make some products appear profitable when they are not.

The financial goal for a profit-making enterprise is to maximize *net profit* (NP), which is T minus OE. To accomplish that, enterprises have to create products that

generate T, make judicious decisions about I, and manage OE in light of T.

Furthermore, the priorities have to be T, I, OE because that fosters growth. This is the opposite of typical management priorities, which focus relentlessly on cost reduction and thereby hinder growth.

In addition to financial measures with conventional names, TA_G has control measures with unconventional names. *Throughput Dollar Days* (TDD) indicates whether work is being shipped on time. *Inventory Dollar Days* (IDD) indicates whether excess inventory is accumulating. TDD and IDD thus steer the manufacturer toward its goal.

At the highest level, *Throughput Accounting for Services* (TA_S) is virtually identical to TA_G . That is, TA_S changes financial measures, and all measures derived from them, plus it changes management priorities. Where TA_G and TA_S differ is in the details because PSTS are the services least like manufacturing.

- Rather than generating T from products, PSTS enterprises generate T from project deliverables and process service levels. Truly variable costs are for things consumed in the production of a service, such as parts used for repairs.
- Rather than investing in plants and materials, PSTS enterprises invest more in skills, intellectual capital, assets, and service production systems. Furthermore, bids and proposals are significant investments.
- Rather than including factory labor in OE, PSTS enterprises use labor from professionals, scientists, and technicians. SG&A costs include labor from partners and principals, whose job it is to sell services engagements.

Just as there is no product cost construct in TA_G , there is no service cost construct in TA_S . OE is simply summed. It's not allocated to services. This avoids the distortions that make some services appear profitable when they are not.

TA_S also has its own control measures with unconventional names. *Project or Process Dollars per Day* (PDD) indicates whether engagements are being completed on time. *Resource Dollars per Day* (RDD) indicates whether excess resources are present. PDD and RDD thus steer the service provider toward its goal.

TA_S creates several benefits for service providers. Management priorities are realigned for profitable growth. Service mix decisions are not distorted by cost allocation. Control measures steer the enterprise toward its goal. Finally, optimization is achieved globally—across the enterprise—rather than locally within a single department or business unit.

Nonstandard TOC Applications

Standard TOC applications are largely consistent across enterprises within a given sector. They include the R, CC, DBR, and TA applications just seen.

In contrast, nonstandard TOC applications vary across enterprises because they have unique requirements. Nevertheless, the underlying TOC principles are still the same, even for nonstandard applications.

- Marketing creates compelling offers.
- Sales then close deals with customers.

By creating offers that address customers' core problems, TOC offers business value that can be much higher than conventional offers heavily based on price.

- Strategy defines the way an enterprise pursues its goal.
- Change then realigns marketing, sales, and production to carry out that strategy.

By creating holistic solutions targeted at constraints, TOC uses leverage to generate large benefits from modest investments.

- Implementation puts TOC applications into practice.
- Technology is a necessary enabler.¹

By using a specific set of steps to achieve adoption and by applying technology prudently, TOC addresses the biggest impediments to implementation.

These nonstandard TOC applications are just as relevant in services enterprises, of course. Details are beyond the scope of what can be covered here, however. See Ricketts (2008, Chapters 8–10) for more information.

How to Cause the Change

What managers know about TOC and what they do with TOC are seldom the same. There are several hurdles. First, when initially becoming aware of TOC, the word “theory” puts some managers off. If they misinterpret “theory” to mean “something that won’t work in practice,” they can’t even get over the first hurdle. On the other hand, if they correctly interpret “theory” to mean “a clear understanding of cause and effect,” they’re off to a good start (Avraham Y. Goldratt Institute, 2008).

As for the second hurdle, TOC is at odds with a lot of conventional management wisdom, as described in previous

sections. When confronted with a demonstration that some conventional wisdom is actually incorrect, most managers react by clinging to it, if not actively defending it. After all, they ask, if so many people believe it, how can it be wrong?

Of course, paradigm constraints like this are the reason why a generation may pass before truly revolutionary ideas in any arena take hold. Therefore, this strong tendency to cling to conventional wisdom is the reason TOC authors and consultants help managers identify core problems, and the faulty assumptions behind them, before introducing the TOC applications that solve those problems (Scheinkopf, 1999).

When an individual manager or a small group of managers gets past the first and second hurdles, there is still a major hurdle ahead. Getting all the other managers, executives, and employees in an enterprise to recognize and accept that conventional wisdom is wrong is arguably the highest hurdle of all. This barrier is, however, precisely what creates sustainable competitive advantage for those enterprises that make the leap to TOC.

Managers are not the only group who can and should be persuaded to adopt TOC. There are also the practitioners, partners, and principals in PSTS. It's not enough for managers to pursue TOC if the people who will be executing the TOC applications aren't convinced as well.

Students, and the professors who educate them, are another vital constituency. Acquainting the next generation of managers is an obvious way to get more TOC adoption, but that requires acquainting the current generation of professors. Those professors often have dual interests in research as well as teaching. Regardless of group, TOC has a specific method for gaining commitment.

Buy-in

The TOC approach to change is called buy-in. Although it sounds counterintuitive, TOC recognizes that the strongest force for change is initial resistance against change. That is, once someone is convinced the situation will improve, there's no longer a reason to resist change, and the commitment to change is stronger than it would have been without this flip in perspective.

Buy-in proceeds in these steps, which must be performed strictly in order (Goldratt, 1999):

1. Agree on the problem.
2. Agree on the direction of the solution.
3. Agree that the solution solves the problem.
4. Agree that the solution will not lead to significant negative effects.
5. Agree on the way to overcome obstacles to implementation.
6. Agree to implement.

Although anyone can follow these steps if sufficiently knowledgeable and motivated, TOC consultants often help clients through these steps because conventional wisdom resists change so strongly. Ironically, those TOC consultants can have just as much difficulty taking their own enterprise through these steps.

How Practitioners Can Get Started

One way to get started with TOC is by studying success stories. They are not hard to find. Many have been published in books and articles. Some can be found on the

Web in blogs. Of course, peers who have successfully implemented TOC are perhaps the most credible source of all.

It's also possible to get started with TOC by hiring TOC consultants. Their breadth of experience often exceeds what peers know because consultants have the additional advantage of having seen what works and what doesn't work. Furthermore, if the consultant brings software assets, then that can aid in implementing TOC.

Certification is another way to get started because it requires completion of formal training as well as exams.² It thus requires demonstrating a level of proficiency above what can be attained on the job or via independent study. This is obviously a path that TOC consultants take.

How Researchers Can Contribute

Researchers also have a role to play in causing change. Field studies, case studies, and simulation studies are all ways to investigate TOC and foster its adoption.

Literature reviews are another potential contribution whose value should not be underestimated.³ When well done, such reviews are used both by practitioners and by other researchers. However, the TOC literature is scattered across several fields and many journals, so the best literature reviews synthesize findings from disparate sources.

For research to have impact outside of academia, it has to be consumable by practitioners and students. This is a challenge because scientific terms and methods that researchers take for granted are alien to non-researchers. Moreover, TOC has its own jargon that can be baffling to newcomers. The result can be publications that are incomprehensible to the audiences who could benefit most.

What Students Should Know

Obviously, students need to know about TOC principles and applications. The more hands-on their education is, the more likely they are to retain what they learn. For instance, simulation games require students to assume roles and play out scenarios based on TOC.

Simulation games are a staple of TOC education, but there's really no substitute for seeing TOC in practice. Plant tours, industry speakers, thesis projects, and internships are well worth considering. Some schools have students work with local firms and conduct a *Thinking Process* (TP) project to identify what to change, what to change to, and how to cause the change. They then present their recommendations to the manager of the firm.

Students also need to know about the TOC buy-in process because students equipped only with a toolkit of TOC principles and applications will run headlong into opposition outside the classroom. Indeed, some enterprises have a latent pool of untapped TOC knowledge because recent graduates in management programs have almost certainly been exposed to TOC during their education. After graduation, however, they wind up in jobs where no one in their management chain is aware of TOC, let alone comprehends it.

Although many graduates have been exposed to TOC_G, few have yet been exposed to TOC_S. That is changing, however. *Service Science, Management, Engineering, and Design* (SSMED) is an academic initiative that involves a broad community spanning academia and services enterprises. SSMED helps academic institutions with curricula and other resources.

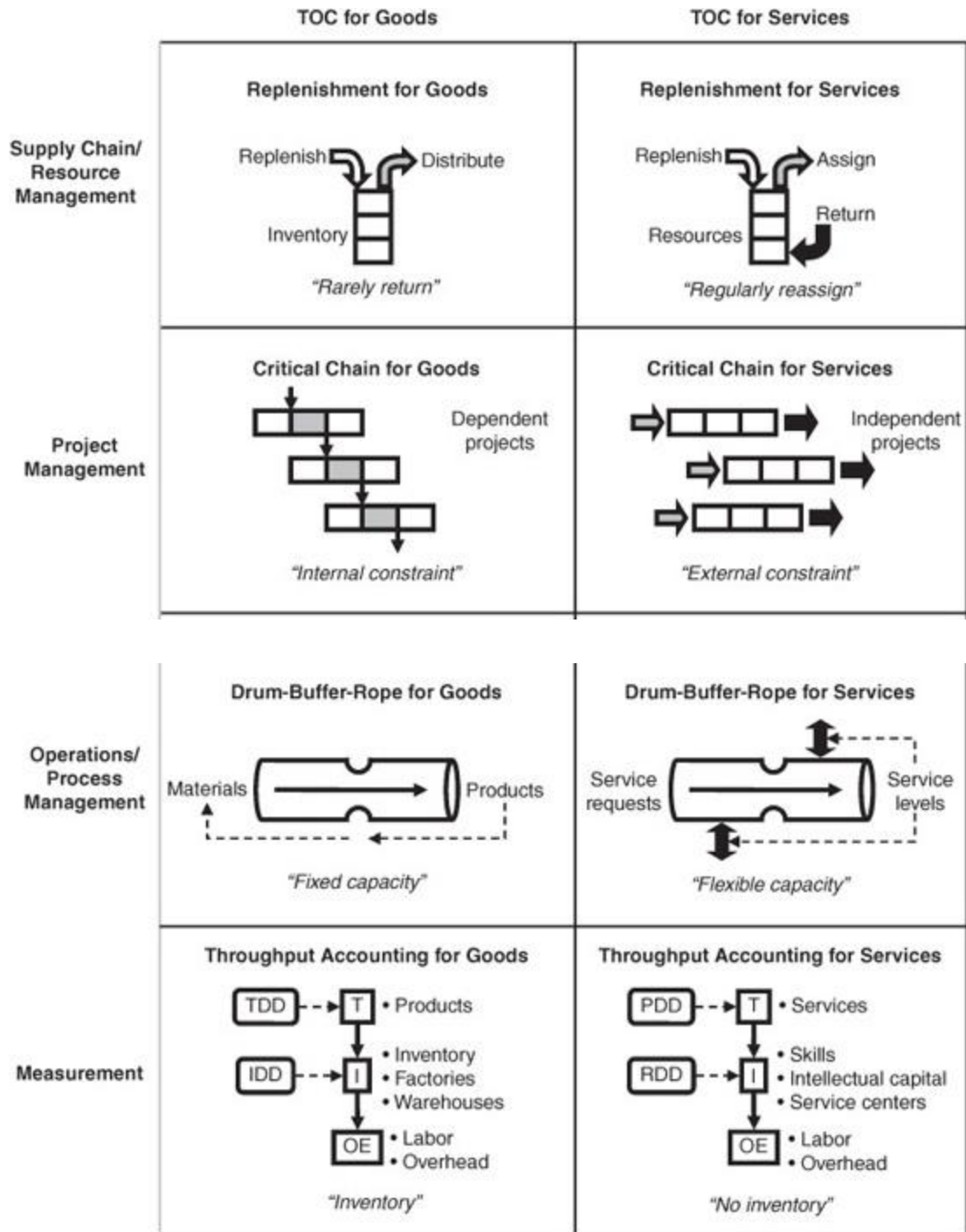


FIGURE 29-1 TOC vignettes.

Summary

The appeal of TOC comes from its sound management principles, plus applications that embody those principles. Here are some examples.

- Drum-Buffer-Rope is based on the *Weakest Link Principle*, which says a system can only produce as much as its constraint will allow.
- Replenishment is based on the *Aggregation Principle*, which says inventory or resources are best buffered centrally because that's where consumption varies least.
- Critical Chain is based on the *Relay Race Principle*, which says work rules (execution) determine on-time project completion far more than the project plan does.
- Throughput Accounting is based on the *Measurement Principle*, which says you have to measure the right things to steer an enterprise toward its goal.
- All these applications are based on the *Pull Principle*, which says the most effective management systems pull work through naturally.

TOC_G and TOC_S are based on the same fundamental TOC principles. Therefore, they are complementary. [Figure 29-1](#) shows TOC vignettes.

- R_G manages inventory that rarely returns once shipped.
R_S manages resources that regularly return for reassignment.
- CC_G manages projects when the enterprise constraint is internal.

CC_S manages projects when the enterprise constraint is external.

- DBR_G manages operations when capacity is relatively fixed.

DBR_S manages operations when capacity is relatively flexible.

- TA_G provides measures when inventory is abundant.

TA_S provides measures when there is no inventory.

PSTS is the services sector most different from manufacturing and distribution where TOC began. Because TOC works in PSTS, where the conditions are extreme, there's a good chance TOC will work in any services business.

References

Avraham Y. Goldratt Institute. 2008. *The Theory of Constraints and Its Thinking Process*. New Haven, CT.

Corbett, T. 1998. *Throughput Accounting*. Great Barrington, MA: North River Press.

Goldratt, E. 1999. *Goldratt Satellite Program Session 6: Achieving Buy-in and Sales*. Broadcast from Brummen, The Netherlands: Goldratt Satellite Program.

Goldratt, E. and Cox, J. 1992. *The Goal: A Process of Ongoing Improvement*. 2nd. rev. ed. Great Barrington, MA: North River Press.

Goldratt, E. M., Schragenheim, E. and Ptak, C. A. 2000. *Necessary But Not Sufficient: A Theory of Constraints Business Novel*. Great Barrington, MA: North River Press.

Mabin, V. and Balderstone, S. 2000. *The World of the Theory of Constraints: A Review of the International Literature*. Boca Raton, FL: St. Lucie Press.

Ricketts, J. A. 2008. *Reaching the Goal: How Managers Improve a Services Business Using Goldratt's Theory of Constraints*. Boston, MA: IBM Press.

Scheinkopf, L. J. 1999. *Thinking for a Change*. Boca Raton, FL: St. Lucie Press.

Spohrer, J. and Kwan, S. K. 2008. *Service science, management, engineering, and design (SSMED): Outline and References*, January. <http://www.ibm.com/developerworks/spaces/ssme>

Umble, M. and Umble, E. J. 2006. "Utilizing buffer management to improve performance in a healthcare environment," *European Journal of Operational Research* 174:1060-1075.

About the Author

John Arthur Ricketts is a distinguished engineer in IBM Corporate Headquarters. As a consulting partner and technical executive, he has dealt with many services management issues, including those faced by clients in their own services businesses. His work in applied analytics led him to become a focal point on *Theory of Constraints* (TOC), and then to delve deeply into its potential for services management. His book, *Reaching the Goal: How Managers*

Improve a Services Business Using Goldratt's Theory of Constraints, was published by IBM Press.

Dr. Ricketts' research and teaching have won awards from the Decision Sciences Institute and the Association to Advance Collegiate Schools of Business, as well as IBM. Prior to joining IBM, he was a professor, manager of applied research, and director of software engineering. Since joining IBM, he has worked on business development, service delivery, professional development, intellectual capital development, and strategic initiatives. His graduate degrees are in management and information systems, with supporting fields in computer science and behavioral science.

CHAPTER 30

Customer Support Services According to TOC¹

Alex Klarman and Richard Klapholz

Introduction—the Need for Change

For several years, *Customer Support Services*² (CS) operations were viewed as an enhancement to product or service sales and a significant revenue generator in itself. Over the years, however, the environment has changed dramatically where CS is providing only a marginal revenue stream at best in many operations.

The purpose of this chapter is to provide a “when to” and “how to” guide for analyzing the problems and designing a practical solution in the area of CS in product organizations. It shows how successful companies accumulate a very large installed base of their products over the years, with various commitments to its users. It can be either a blessing or a curse.

Good CS can be a significant asset, creating opportunities for repeat sales at low cost and effort. However, problems in this field can require ever-growing resources and steadily diminishing returns, as well as possibly jeopardizing future business relationships with clients and users.

The various domains of CS are presented, including the warranty, the nature of service contracts, the impact of CS on the revenues and the expenses of the firms, and the

resulting impact on its bottom line. An analysis of the pertinent problems in all of these areas is addressed as well. The direction of the solution and the major solution components are described with respect to how they solve the core problems and therefore eliminate the limitations of the traditional approach to CS. Additional supporting actions required to provide a complete solution are given. Implementation and day-to-day management issues are discussed.

What Is Customer Support (Also Known as Technical Support)?

CS is a vast area of modern economy, and there is hardly a product or service that does not necessitate the use of it. Be it a cell phone, an electric can opener, a notebook computer, a TV cable service, a food processor, a new car, a new computer game, or even an old mattress (but one with a lifetime guarantee)—they all sometimes need external assistance with their proper installation, use, maintenance or repair, and finally, proper disposal.

The growing complexity of modern products, the richness of their features and the mind-boggling speed of technological advancement all make usage of new products (or services) a challenge to all but the few technically gifted (or simply very young). That, when coupled with ever shorter life cycles of most products, makes a thorough knowledge of all of a product's (or service's) features truly "Mission: Impossible," and makes external help an essential part of our daily life. However, the fact that it is such a daily affair does not make it any easier. Quite often, it is an ordeal we have to endure if we are to enjoy the benefits our modern world has to offer.

There is hardly a person who has not witnessed first hand both the best and the worst of technology; from a help-desk person who has revived a "dead" computer with a few

simple instructions to a Kafkaesque ordeal, often involving automated replay systems.

In this chapter, however, we will limit ourselves to CS of industrial equipment and software, which is also called technical (or tech, for short) support.

So, what's this CS we talk about?

Of the many definitions one can come by when conducting a basic search on the Web, two main characteristics stand out: One deals with making a proper (and cost-effective) use of the product or service, while the other focuses on customer satisfaction. For example:

- The range of services designed to assist customers in making cost-effective and correct use of products or services. It may include assistance in planning, installation, training, troubleshooting, maintenance, upgrading, and disposal of the product (or service), as defined by www.BusinessDictionary.com.
- According to Turban (2002), "*Customer service is a series of activities designed to enhance the level of customer satisfaction—that is, the feeling that a product or service has met the customer expectation.*" For the purposes of this work, we will use the following definition, which combines both elements of the proper usage as well as that of customer satisfaction:

Range of services designed to enhance the level of customer satisfaction—that is, the feeling that a product or service has met customers' expectations. It is achieved by assisting the customers in making cost-effective and correct use of product or service. It may include assistance in planning, installation, training, troubleshooting, maintenance, upgrading and disposal of the product (or service).

At once, this definition outlines the critical importance of customers' *expectations* in achieving the goal of providing a good CS. It also shows the wide range of CS activities, which can spread over the entire life span of a product.

However, it clearly shows that a significant part of CS activity comes quite early in the life cycle of the product; during the phases of planning its usage, the installation of the equipment (or service) and its initial use, together with the training of the staff in the proper procedures of exploiting it.

Usually, the first period of product (or service) use provides CS free of charge to the users; it is the *warranty* period. It is often then, when mundane reality, with all its problems, clashes with customers' illustrious expectations. True, quite often these expectations are derived from what the marketing and sales personnel have communicated—be it explicitly or implicitly—to the future clients. The resulting perception stemming from the comparison between the expected and what was actually delivered creates a lasting impression, which will (almost) forever influence—for good or for bad—the relationship between a particular customer and the product or service provider.

Thus, the quality of CS one receives may be a crucial element in future business decisions regarding repeat buy of a product or a service. Nowhere do these decisions have more impact than in the business of industrial equipment.

In this work, we will try to show how TOC and its application can significantly contribute to successful CS, and thus to a better, more successful business.

Steady Erosion of Income in the CS Area

Three main processes, which have steadily advanced over the course of the last few decades, contribute heavily to the problems the CS area faces today:

- In many key industries (printing, metalworking, textile production, and microelectronics, to name a few) the selling price of the equipment has decreased considerably. That, of course, made the equipment within the reach of many potential buyers who could not afford it earlier. The resulting growth of sales volume was huge. When coupled with the growing competition as a result of developing countries joining the ranks of the established production centers of the West and East, the economies of scale due to much larger markets, and the advent of electronics replacing mechanical or optical solutions, they enabled production of more and more sophisticated equipment at prices that continuously decreased.
- At the same time, the growing complexity of the equipment put even greater and more varied demands on CS providers. Equipment, which was once the provenance of only the largest organizations with their own engineering and maintenance resources, now became available to much smaller firms. Once the equipment operates at these smaller firms, in which the division of labor between day-to-day operations and the technical support function does not really exist, it may create a problem. There is an expectation that somehow the equipment will always be functionally available, without the need of creating (and paying for) a specialized internal maintenance unit to make it happen. At such firms, the dependence on the goods' producer to provide the technical support is crucial.
- There is an additional effect having its impact; the life cycle of equipment undergoing a continuous reduction, resulting in ever-shorter periods of time between the appearance of a particular piece of equipment and the arrival of its successor. Even in the short period of time between equipment generations, there is a steady

stream of improvements, upgrades, changes, and additions to it. For the equipment producers it meant, if they are not to continuously increase their engineering ranks, they would have a great and growing strain on the engineering resources. As a result, quite often the products arrive in the marketplace before the full development process is completed; the rigorous testing procedures in various scenarios and multitude of operational environments, a true *must* at products directed to mass markets, is often cut short. That, in turn, brings about significant “growing pains” to the users of the products. The part of the equipment producers’ organization, which will have to deal with the problems stemming from this is, of course, the CS unit. Many firms keep the older equipment after newer versions are available, forcing the CS staff to master the maintenance of a number of versions of its equipment, long after many have been replaced.

However, with all the dynamic growth it underwent, the modern industrial equipment market does not have the huge size of the true mass market (like that of cell phones, cars, digital cameras, or laptops). In contrast to the true mass production markets, the industrial equipment offer to the market is characterized by a wide variety of products and features, relatively low volume, and almost always high complexity. These characteristics do not allow for the economies of scale needed to bring its development process, as well as its production procedures, to deliver truly fail-proof products.

On top of these phenomena, which characterize a large part of the products market, there is a rather peculiar system for setting the price of the CS provided to the users of a particular product or service. Usually, after the conclusion of the warranty (guarantee) period, the service is

provided at a fixed annual cost, which is some agreed-upon percentage of the selling price of the given product. Be it 1 percent or 20 percent, in exchange for it, the supplier is bound to supply a service, which grants the users the full functionality and service response they are expecting.

As with other parts of any business organization, CS is striving to contribute its share to the bottom-line results of the firm.

For many years in so many companies, the revenues of the CS organization were a major source of income. Moreover, as the annual CS contracts were almost automatically renewed year in year out, it was a source of very steady income, independent of the vagaries of the sales efforts. However, let us not make a mistake; the income derived was as a rule the result of hard work of top-notch professionals. At the base of it was an expert knowledge, a result of years of hands-on experience, which enabled it to happen.

As we have mentioned previously, the growing competition has brought about, with a plethora of other effects, a constant process of decline of the selling price of the equipment. That, in turn, has reduced the income derived from providing technical support for the installed equipment. Moreover, today the price, which is the basis for comparison between different suppliers of equipment, is the *total cost of ownership (TCO)*, which takes into account not only the selling price, but also the expenses involved in keeping that equipment fully operational.

Fierce competition imposes pressure on equipment price and thus on the price of technical support, which leads to a steady erosion of service revenues. The net result is the fast transformation of what used to be a very profitable part of the business into a problematic one. Its impact on the overall profitability of the equipment business is turning from very positive to much less so.

The interplay of these phenomena can be described as the following cause-and-effect (*Current Reality Tree* [CRT] in TOC terminology) as diagrammed in [Fig. 30-1](#).

One can easily envision what will happen over time, if the Operating Expenses involved in providing service is fixed (if not growing over time), while the revenue it generates only goes down, as in [Fig. 30-2](#).

In the business world, if a product or service shifts from a profit-generating operation into a losing proposition, it must be corrected, abandoned, or replaced by a better one. However, what if such a loss-generating operation happens to be the key to customer satisfaction and loyalty? In many cases, CS is the main driver in future sales to existing customers. In addition, the satisfied existing customers provide references to potential customers.

However, this deterioration of profits, in turn, leads equipment manufacturers to realize that they already face (or, in no time they certainly will) the following dilemma, with no simple solution in sight, as we see in [Fig. 30-3](#).

The Warranty Trap

The initial period of time during which CS is provided, generally free of charge, to the user of the equipment is usually called the warranty period. Warranty is the assurance of the seller to the purchaser that the goods will function properly and shall be as represented. If not, it will be replaced or repaired. It constitutes a part of the purchase contract and must be fulfilled to keep the contract in force.

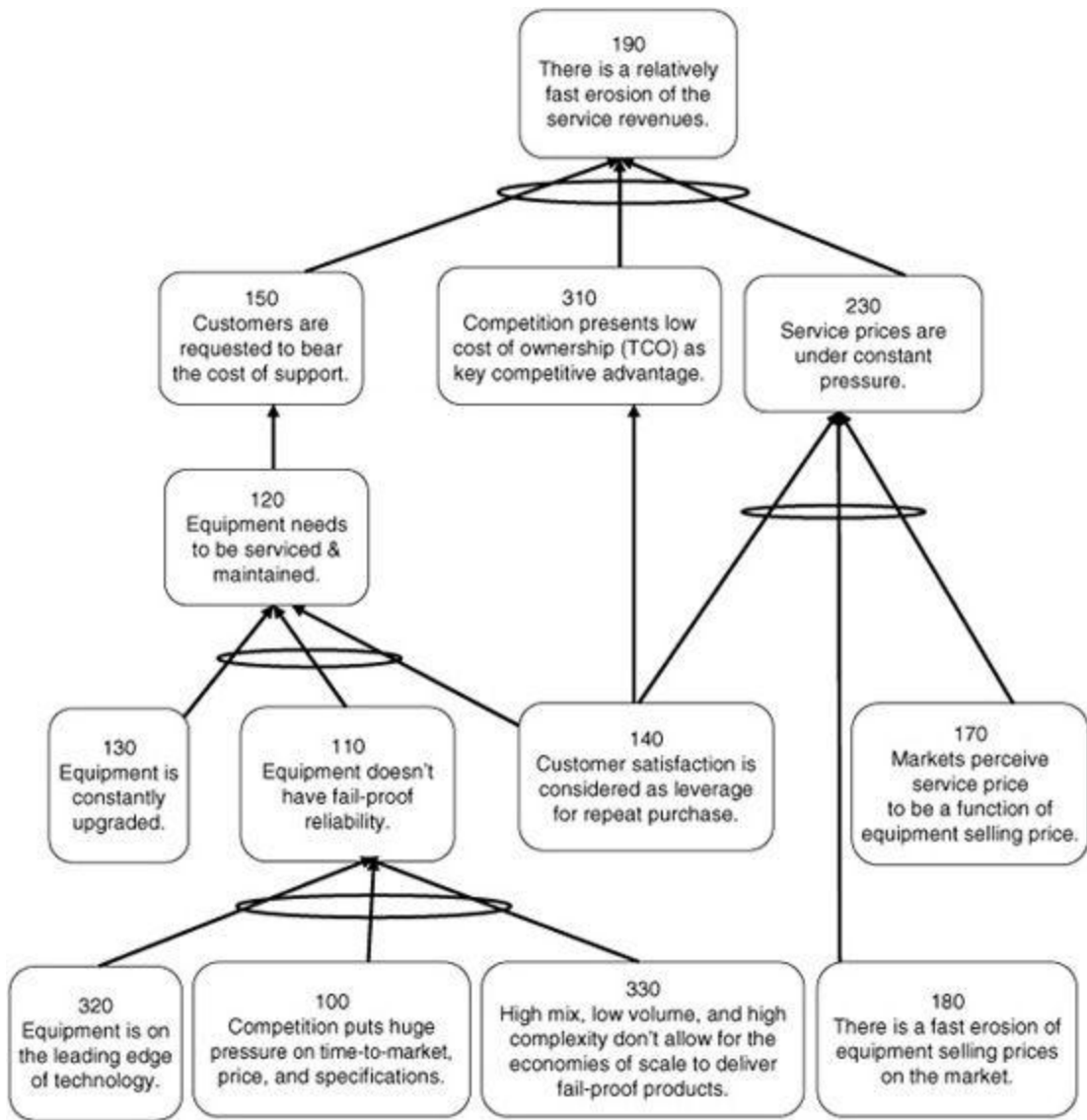


FIGURE 30-1 CRT of customer support. (Source: Modified from Klapholz and Klarman, 2009, 13.)

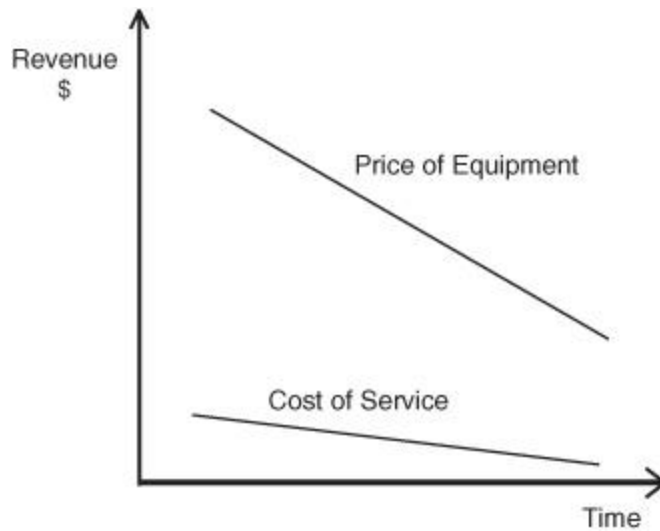


FIGURE 30-2 The dwindling CS revenue. (*Source:* Klapholz and Klarman, 2009, 24.)

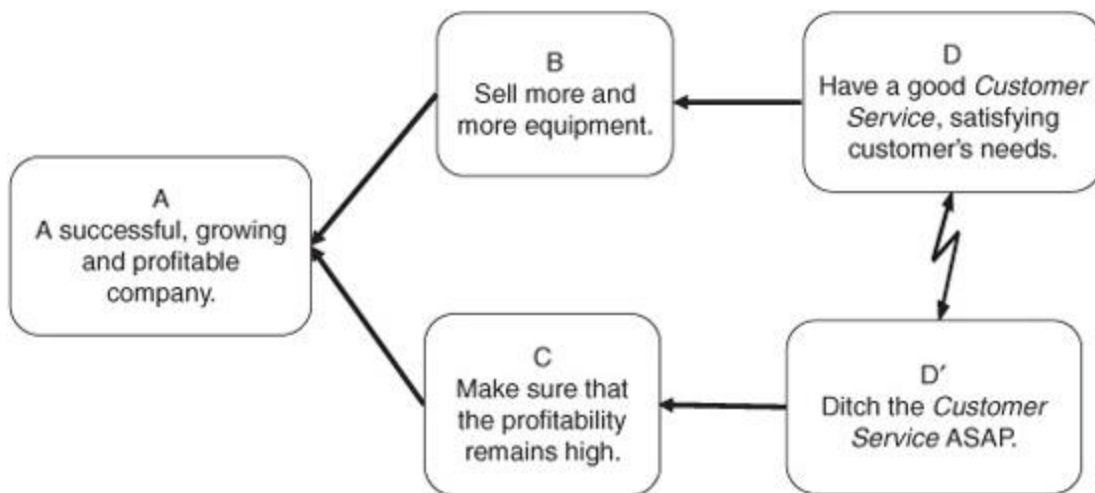


FIGURE 30-3 The dilemma of CS. (*Source:* Klapholz and Klarman, 2009, 19.)

From the viewpoint of the customers, it constitutes a critical component of the value they expect to derive from the purchased equipment or service. However, for the customer service organization, which provides this service (usually) without a charge to the customer for a period of time, it can be a tricky business. Although it is customary to

present this in the firm's accounting system as warranty revenues with its associated expenses, and to talk about warranty profitability (or loss thereof), it is not really managed that way from a business point of view.

Warranty revenues are usually a time-proportionate, fixed percentage allocation of equipment sales revenues to CS. It does not reflect the value of the service agreements that will replace it once the warranty expires; it is just an arbitrary allocation—an allocation, not a market price, tested through the market's competitive mechanism.

Warranty expenses are usually “buried” in the overall expenses of CS. It typically covers not just service, but also the installation and training of the user. In addition, the timing of the beginning of the warranty period is often vague, and frequently tends to start long after the shipment of the equipment; sometimes after its installation, often after formal client acceptance. One can only imagine, the faster the product is rushed to the market the bigger the burden of warranty expenses, without any comparable growth in the percentage of revenue allocated to CS.

In addition, one of the “goodies” sales people use to “sweeten” the sales deal is the extension of the warranty period, either free of charge or at minimal cost to the client. This additional burden on CS is rarely (if ever) factored into the business picture.

No wonder that warranty expenses, which in the past were a minor irritation in an overall positive picture, are seen today as a major problem; a problem that casts a dark shadow on the overall CS business situation, which is fast moving from rosy to a very worrisome one.

As we can see in [Fig. 30-4](#), warranty revenues distort (see the core problem in 500) rather than assist the ability to analyze, forecast, or plan warranty's impact on the overall business of CS.

When we add what happens with the warranty to the problems CS already faces (lower prices, greater strain on

the resources, shorter product life cycle), the overall business picture of CS becomes even more grim (see [Fig. 30-5](#)).

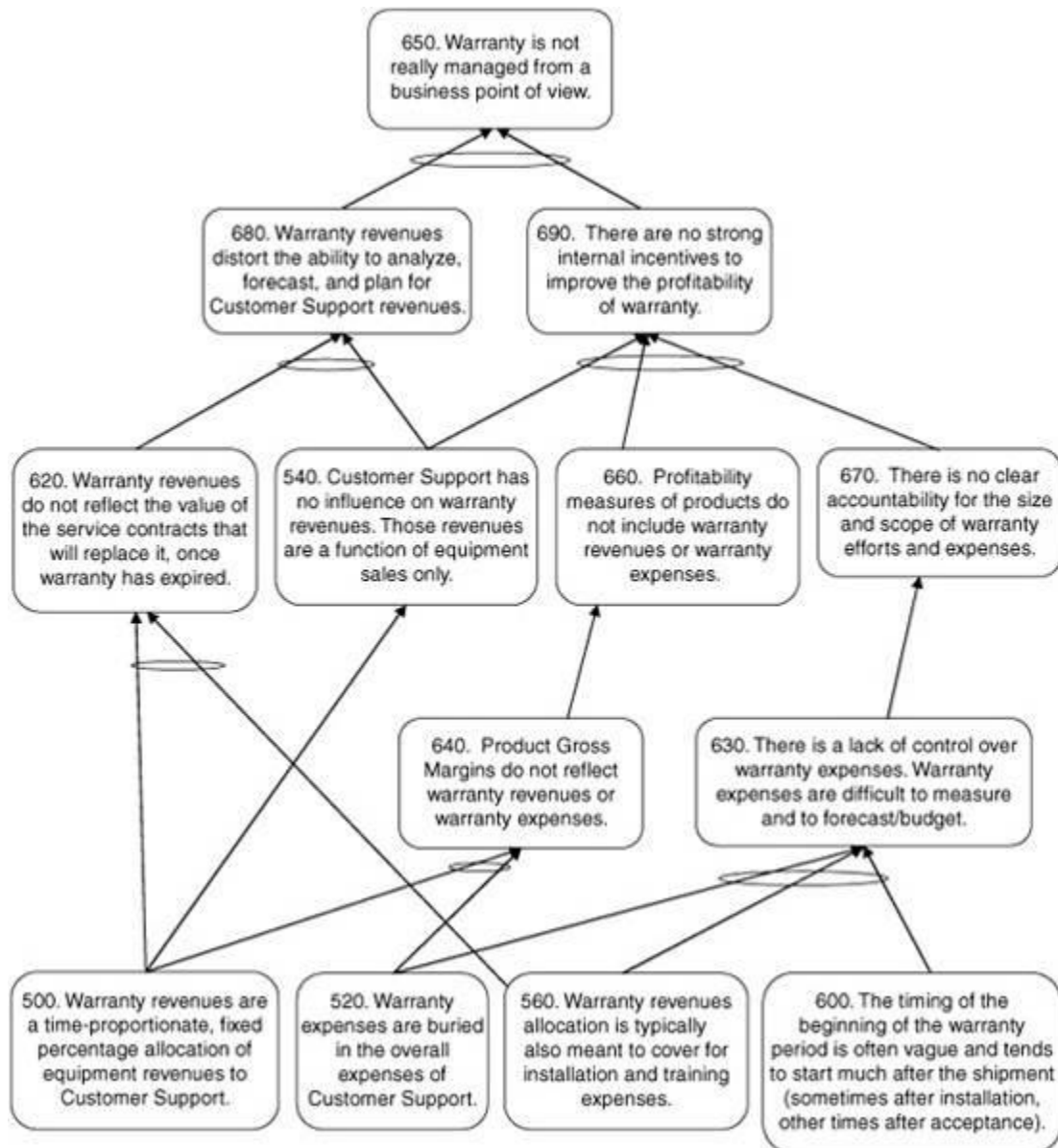


FIGURE 30-4 Warranty CRT. (Source: Klapholz and Klarman, 2009, 47.)

So, why not dump CS altogether? Why should the producer take this burden in the first place? True, you hardly can sell anything today without providing proper warranty

and service (is watermelon the only thing sold without one?), but why not let the market mechanisms take care of that? The problems lie in the equipment producers' realization that CS is necessary due to its strategic impact on the revenue of the firm, both in the present as well as in the future, as seen in [Fig. 30-6](#).

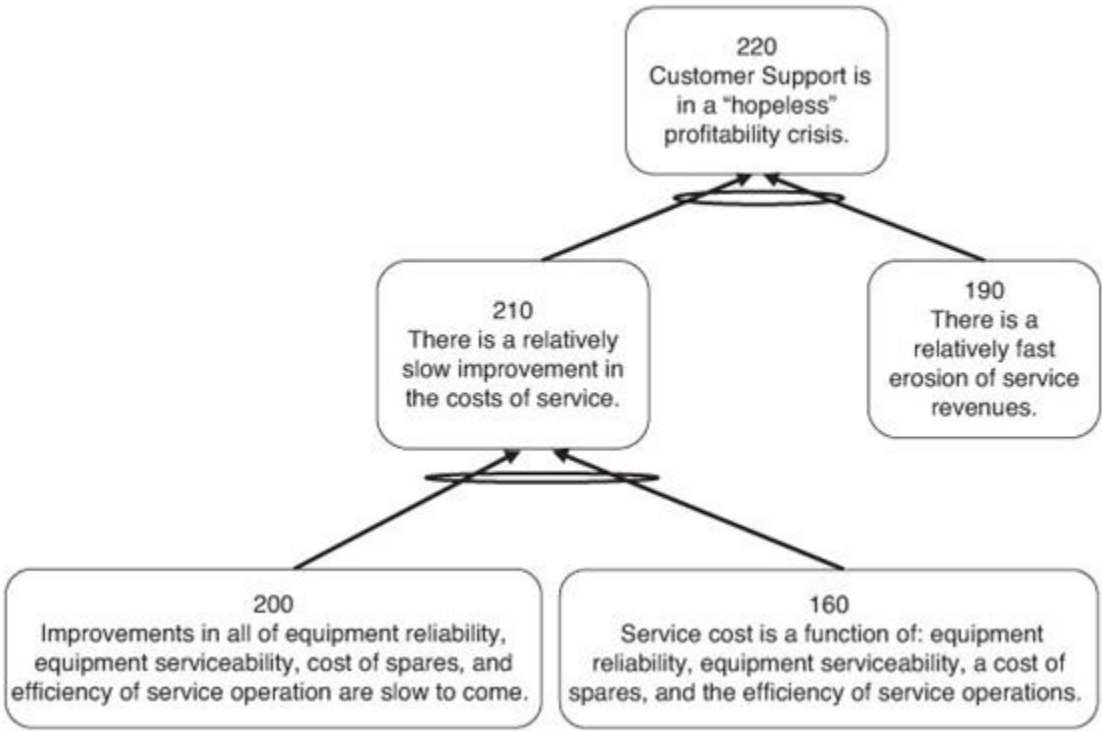


FIGURE 30-5 The bleak outlook of CS. (Source: Klapholz and Klarman, 2009, 14.)

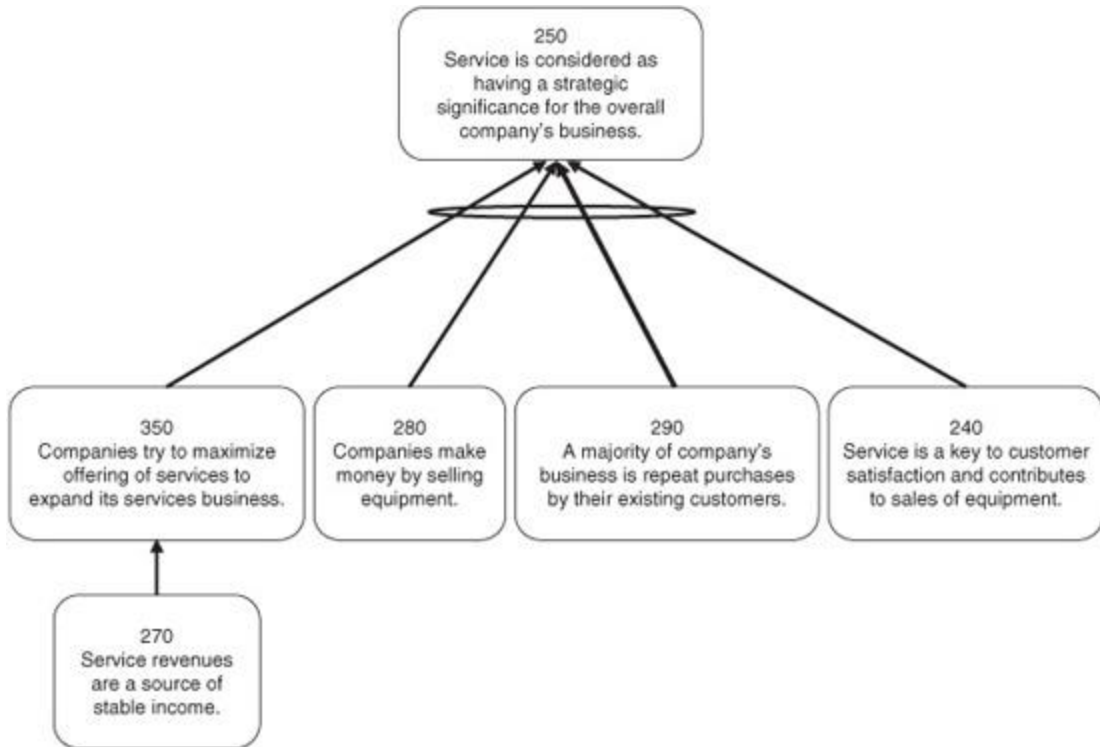


FIGURE 30-6 The business impact of CS. (Source: Klapholz and Klarman, 2009, 14.)

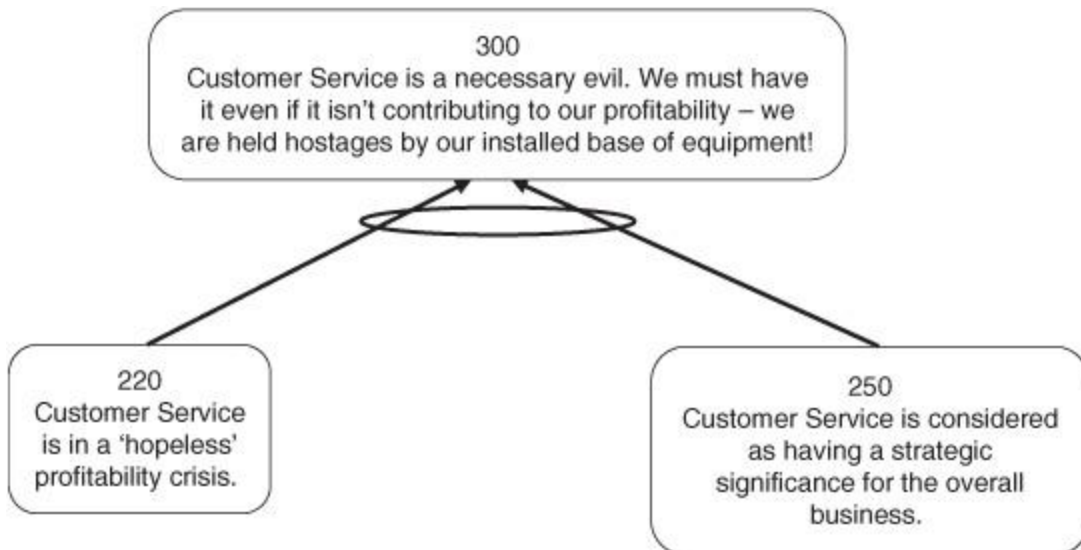


FIGURE 30-7 CS as a hostage. (Source: Klapholz and Klarman, 2009, 15.)

The bottom line of this situation is an unattainable one; companies in the equipment business are held hostages of their installed base of equipment as in [Fig. 30-7](#).

It is far away from the common wisdom of equipment producers, which for years have seen CS as a reliable “cash cow” of their organizations, immune to the inherent capriciousness of the markets.

What to Change

The basic approach of TOC to problem resolution requires us to start with identification of the core problems—the core drivers behind the rise of the multitude of undesirable effects we endure in the organization.

In order to make sure that the reality picture we portray here, dark as it may be, is an exhaustive one, we must go one step further; let us have a look at the day-to-day relationships between CS and its customers. We should relate to what CS personnel often call the “abuse” of its services by the customers.

In spite of the mantra of any business organization, “The client is always right,” nobody knows better than CS that it is not always the case. The CS operational *motto* is rather, “The client is not always right, but he always is a client.” The CS contracts are similar to insurance contracts, promising the insured that when in need, the insurer will provide the technical expertise needed to bring the situation back to normal. Of course, like in any binding contract, there are plenty of *details outlining what service clients are entitled to* (called *Service Level Agreements*, or SLAs for short). Nevertheless, the very structure of these contracts leads to a very inefficient operation—from the point of view of the service provider—of the service organization. The most common form of equipment service contract is that called a “full service contract,” which provides full coverage,

without regard to the effort or expenses involved in providing it. However, unlike in insurance, there are no “deductibles” and “copayments,” and there are no discounts to customers with no (or less) claims. It creates an inherently abusive situation, because:

- The customers have every incentive to call for help, even if the problem is minor and can be solved by them. “The customer is always right” is the slogan that rules here.
- So what we face here is a situation in which the revenues (a fixed fraction of the purchase price of the equipment) are controlled by the cutthroat competition between producers and generally decrease with time, while the expenses are an open-ended and mostly increasing proposition.

All that means only one thing: If CS is to refrain from becoming a bottomless pit, relentlessly swallowing the revenue generated in other parts of the organization, it must reinvent itself. The improvement methodology of TOC provides a proven path to achieve this ambitious feat.

The problem is a tremendous one, as it is clear that both preconditions of the dilemma ([Fig. 30-3](#)) are truly mutually exclusive:

Having good CS; satisfying customers’ needs is diametrically opposed to the demand: Ditching the CS ASAP.

What to Change to

The TOC approach calls for a clear visualization of the basic conflict, which prevents the resolution of the core problem.

If we realize that good CS is a must, our question isn't whether or not we want to have it, but rather what is needed to change for it to contribute what it should to the firm's overall profitability.

That means that the D' want of the "cloud" in Fig. 30.3, stating "Ditch the CS ASAP" is unacceptable. This, in turn, translates into the question, "How do we make sure profitability remains high (need C), while providing good CS, and satisfying customers' needs (want D)?"

The Evaporating Cloud in [Fig. 30-8](#) shows the inherent conflict between what CS sees as waste (providing *unnecessary services*), and what its clients regard as their paid-for right, almost a birthright (*all services* are provided upon request). It stems from both sides of the profitability equation (as the objective A is to increase profitability); increasing the efficiency and effectiveness (need B) by lowering the unnecessary expenses (want D), but still preserving the revenues from service contracts (need C) by providing services clients want.

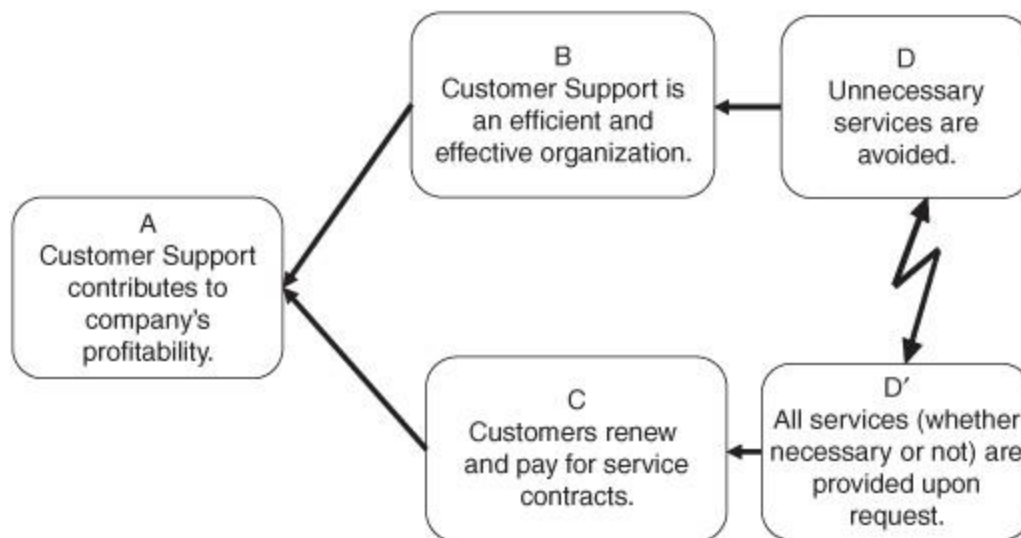


FIGURE 30-8 The dilemma of CS: what service to provide? (Source: Klapholz and Klarman, 2009, 74.)

So how is one capable of providing all services needed by the customer, while refraining from providing what is unnecessary? What change in the existing service providing arrangement could be of value in both the eyes of the service users *and* the service providers at the same time? How can both sides benefit from it? Let's have a look at the assumptions underlying this chronic dilemma in [Fig. 30-8](#).

A—B

AB1: Efficiency and effectiveness bolsters CS contribution to the company's profitability.

AB2: Efficiency saves time and expenses, thus adding to CS contribution to the company's profitability.

AB3: Effectiveness saves time and expenses by eliminating unnecessary activities, thus contributing to the company's profitability.

AB4: Effectiveness increases Throughput, thus contributing to the company's profitability.

A—C

AC1: A significant part of a company's revenues comes from the renewed service contracts.

AC2: Providing free services hurts the company's profitability.

AC3: Paid-for customer service has a significant impact on the company's profitability.

B—D

BD1: Providing unnecessary services hurts effectiveness.

BD2: Unnecessary services compete for the same resources, which may be busy providing truly necessary service.

BD3: Unnecessary services may constitute a significant part of the CS workload.

BD4: Unnecessary services may constitute a significant part of the CS expenses.

BD5: Unnecessary services do not contribute to CS revenue.

BD6: There is an ability of CS to distinguish between the necessary and the unnecessary services.

C—D'

CD'1: For the customers, service of their equipment is a necessity.

CD'2: It is almost impossible for customers to have an in-house expert covering at all times all of their equipment technology needs.

CD'3: In the long run, nobody will renew a service contract if CS is not capable of providing the necessary assistance at a time of need.

CD'4: In the long run, nobody will pay for a service if its provider is not able to provide the necessary assistance at a time of need.

CD'5: Service contracts do not discriminate between necessary and unnecessary services.

CD'6: Clients sometimes request what turns out to be an unnecessary service.

CD'7: Clients sometimes are unable to distinguish between necessary and unnecessary services.

D—D'

DD'1: Not all requested services are truly necessary.

DD'2: There is no ability on the part of the customer to distinguish between the necessary and the unnecessary services at all times.

DD'3: There is no need on the part of the customer to distinguish between the necessary and the unnecessary services at all times.

The key assumptions we would like to challenge lay behind the D-D' conflict arrow; it is the lack of either the ability or the need (or the will) on the part of the clients to differentiate between what is truly needed and what is not (assumptions DD' 2 and 3). It assumes that quite a significant part of the services customers request is superfluous and not really necessary. CS knows this distinction, but is prevented from acting upon it for fear of losing future contracts, while the clients are quite oblivious to it.

However, what if we challenge these assumptions? What if we create a reality in which this distinction is as clear to the customer as it is to CS? Moreover, if we succeed in designing an environment in which the interests of both sides coincide, instead of colliding, we may have a solution to our problem, one that stands a good chance of success. This would indeed be a breakthrough injection.

Rarely can a complex problem be solved with just one bold stroke (Alexander the Great and the Gordian knot notwithstanding); we would like to present some of the main changes needed to restore CS contribution to a firm's overall profitability.

Differential Pricing

The very first order of business in resolving our problem is the mapping of unnecessary services (according to the CS personnel) that customers demand. It stems from the BD6 assumption, stating *there is an ability of CS to distinguish between the needed and the unnecessary services*. It is a valid assumption, which we do not challenge. As a rule,

most events deemed unnecessary relate to unscheduled events, not to the routine (and planned ahead of time) service visits. Even when viewed as urgent, even emergency events, a good part of them could be resolved without creating an undue load on the service providers. When scrutinizing CS activity, one can categorize their activity according to the effort or the expertise level needed for their resolution. Moving from the most common and simple to the most complex, they can be listed as follows:

1. Problems the customer can easily resolve.
2. Problems that the Response Center (or Call Center) can fully resolve with the clients.
3. Problems that the Response Center can diagnose, but will still necessitate the on-site arrival of the *Field Service Engineer* (FSE) to repair.
4. Problems that the Response Center has difficulty diagnosing, which necessitates the on-site presence of the FSE for diagnosis.

It is often the feeling of the CS personnel that if the first two types were handled properly by the customers (i.e., as described previously), instead of the hasty dispatch of an FSE to the site, a large portion of the waste could be prevented. If only we could find a way to bring customers to try to resolve problems by themselves first, without turning to CS, or to cooperate better with the Response Center instead of calling for the prompt arrival of an FSE, it would change the business picture radically—and for the better.

As we have said previously, the idea is to create common interest between the customer and the service provider to resolve the problem quickly and effectively. As long as there

is only one business model—the current one—it will not work.

However, what if we provide our customers with a different business model, one that creates incentives to reduce service calls, to make an effort to resolve problems by themselves, to minimize the usage of the service only to the “must” cases, instead of the current “why not?”

The elements of the proposed solution are as follows:

Instead of the current one standard “unlimited” support service for a fixed fee, we can offer our clients a number of differentially priced options. We should price them in a way that will reward them for minimizing demanding unnecessary services, thus creating a truly win-win solution. The following are examples of the range of service options.

The Array of Service Offerings

Basic Services

This is the basic building block of all the service program options and consists mainly of remote services. Some customers who are currently on *Time & Materials* (T&M) only (i.e., they pay per call time of the FSE and per parts/material every time there is an occurrence), may be attracted to this option. Basic Services include:

- Phone support 9:00 AM to 6:00 PM, Monday through Friday, with a two-hour maximum response time.
- Remote access support.
- Remote application support.
- Software upgrades (including implementation of upgrade).

- System review or audit once per year (to ensure that there is no systematic degradation of certain equipment for customers who will take only the basic services).
- CS publications (updates for user manuals, reference guides, quick reference guides, training materials).

Extended Basic Services

Basic Services are a prerequisite.

- Extended hours phone support: 7:00 AM to 11:00 PM, Monday through Friday and 10:00 AM to 8:00 PM weekends, with a one-hour response time.

Limited FSE Visits

Basic Services are a prerequisite.

Consists of on-site visits, including FSE labor and travel costs. Does not include parts.

- Five on-site service visits per product per year, 9:00 AM to 6:00 PM, Monday through Friday, with a one-business-day response time. Visits include preventive maintenance visits (according to product policy) and one system audit per year.
- Right for additional visits at a predetermined fixed price (independent of time and travel; *No-Questions-Asked* [NQA] policy adopted for repeat calls).

Extended FSE Visits

Basic Services are a prerequisite.

Consists of on-site visits, including engineer labor and travel costs. Does not include parts.

- Five on-site service visits per product, 9:00 AM to 6:00 PM, Monday through Friday and 10:00 AM to 8:00 PM during weekends, with a one-business-day response time. Visits include preventive maintenance visits (according to product policy) and one system audit per year.
- Right for additional visits at a predetermined fixed price (independent of time and travel; NQA policy adopted for repeat calls).

Complementing FSE Visits

- Coverage for unlimited visits (in addition to the five visits covered by Limited FSE Visits).

Complementing Extended FSE Visits

- Coverage for unlimited visits (in addition to visits covered by the Extended FSE Visits).

Parts Services

Services must be offered by a Certified FSE (whether from customer, third-party service provider, or company's CS).

- Hardware upgrades for *Field Change Orders* (FCOs).
- Spare parts.

Important Notes

1. All service programs listed previously are presented per product.
2. All product types at the same site are covered by the same service programs.
3. *Value-added services (VAS), such as training or advanced application support*, are not included in the service programs.
4. Pricing will be such that Basic Services, Limited FSE visits, and parts services are the most attractive choices. Extended and Complementing Services will deliberately be priced “out of range” but still made available.

Presenting the market with a range of options, priced in a way that minimizes current “abuse” of services is definitely a big step in the right direction. As we have learned from personal communications of CS managers, it alone can decrease up to half the expenses of the CS department. However, that is far from all that can be done to improve the CS contribution to the company’s profitability.

At least four additional areas present significant potential for improvement (if improvement is defined as either a decrease of the operational expenses or an increase in the Throughput of the CS system). These areas are listed as other service offerings.

Other Service Offerings

Value-Added Services

These are knowledge-driven, high-end, high-margin activities, which enable the equipment user to derive much higher value from its use. Usually, companies turn to external consultants to provide this type of expertise. Often it has to do with better, smoother work-flow organization, better physical arrangement of the machines, and improving the interaction between various departments—all in order to improve the client's positioning in the market. Often it necessitates an in-depth understanding of the client's operation in order to correctly identify the constraints of the system, or more efficiently exploit them, or—and that is (according to our experience) the most common case—better subordinate the entire system to its constraint. As equipment makers, CS departments are often populated by people perfectly suited to perform such tasks; it positions the equipment maker as a better business partner, increasing the chances of future purchases. Furthermore, it transforms CS from its current “break and fix” mode to a consulting-like entity. And of course, such an activity can be amply remunerated, significantly higher than the standard service fees.

Launching of Expert Systems

Quite often large chunks of the expertise required to resolve customers' problems effectively and efficiently are not properly documented and readily available to the technical staff. Usually it resides in the memory of the service providers, and that is one of the main reasons CS is viewed as more an art form than a science. It is just one of the many facets of the problems involved in the organizational

knowledge preservation and management. If, however, the organization makes the necessary effort to build a system to identify, assemble, create, catalog, represent, distribute, and enable adoption of the insights and experiences of its experts, the benefits can be huge to both CS and its clients. Such systems are called Expert Systems, and as such, they make the insights and experiences readily available to everybody in the service organization. They comprise the assembled knowledge, either embodied in individuals or embedded in organizational processes. The potential of having all that expertise readily available, without the need to experience a lifetime of CS work firsthand, can turn even a beginner into a valuable worker, almost from his or her first days on the job. Using a computer expert system to assist CS personnel can decrease even further the need to perform on-site visits. When coupled with good remote diagnostic systems, an expert system has the potential to improve service to clients significantly, while considerably decreasing the costs involved in providing it.

Third-Party Maintenance (or TPM)

Third party maintenance (TPM) is a name given to outsourcing of the CS activities to an external entity, which is capable of providing it in exchange for a price, which is lower than what it currently is when provided internally. We are not dissecting the benefits and dangers of outsourcing here; it is enough to say that it is a move to be considered carefully, as problems in this area can endanger future purchases. Quite often, the third party will require logistical support (like original spares and materials), knowledge transfer (upgrades and changes), and even personnel assignment in resolution of the most complicated cases. Nonetheless, it can considerably offload some constrained resources or create support availability in regions devoid of

it, further increasing sales potential. Sometimes the outsourced work is performed by companies specializing in CS, the so-called *Multi-Vendor Service Providers*. These are companies that provide technical support to a wide range of equipment, produced by a variety of producers. Usually such companies have good presence in wide regions and due to their efficient usage of resources and very low overhead, they can provide their services at very competitive prices.

Installations, Implementations, and Projects

If there is an area in which TOC applications can drastically improve the performance of the CS staff, it is at the first stage of its involvement in the service of the equipment, namely at installation of the equipment or (at the so-called “turn-key” deals) implementations. Options such as bringing the equipment to the status of “up and running” is a multistage activity, usually designed as a project. Even the installation of a simple system is comprised of, at least, unpacking, installation of separate units, their integration, customer training on its operation and day-to-day maintenance, and performance of the acceptance tests. As it goes in projects, it usually takes longer than planned, delaying the start of the warranty period and preventing the team involved from moving to their next commitments. The TOC *Critical Chain Project Management* (CCPM) methodology provides a much better way to deal with the inherent uncertainty characterizing projects while significantly lowering the risk of exceeding the plan’s confines.

An additional domain that can be addressed in order to improve the integration between CS and the entire company has to do with problems stemming from the current arrangement regarding the warranty.

Instead of allocating a fixed, time-proportionate amount from the sales to the service revenues as warranty revenues, a different method is recommended:

1. When a product is sold, some of the product revenues will be deferred-till the end of the warranty period. Those product revenues will accrue during the warranty period on a periodical (say, quarterly) basis. The amount to defer and the length of the warranty are 100 percent business decisions that are made by the product business entity. Of course, the longer the period of warranty, the slower the income accrues to CS.
2. CS will charge the product business entity a “readiness expense” for that product (a rather small amount that covers CS infrastructure expenses such as Response Center, Logistics, etc.) and a fixed amount per every warranty event. CS no longer receives any fixed warranty revenues. The warranty becomes an expense that will be charged to the product business entity on a quarterly basis. The fixed amount per event will be agreed to with the product business entity at the product launch or during the budgeting process. As with every “transfer price,” which is arbitrarily set between two sister units within the same company, one should devote utmost care in setting it. For example, it should be structured in such a way that it will not push one of the units involved to prefer interaction with an external entity rather than the sister unit.

As mentioned already, in this approach there is no longer such a thing as “warranty revenues” as a subsection of CS revenues. Product revenues remain product revenues. This makes sense as, after all, customers refer to what they pay

as an amount paid for the product and all that comes with it: the company name and reputation, the company expertise, the company R&D backing, and, of course, installation, training, and warranty. Why split the revenues at all?

At the same time, warranty expenses are mainly a function of service efficiency, the product quality, and the terms and conditions of the warranty (length in time, limited or full coverage, etc.). While CS determines the first one, the latter two have nothing to do with it. These are determined solely by the company's business entity, which designs the product and ensures its quality at the exit of the manufacturing gate. Now, all the warranty expenses are kept as expenses belonging to the product business entity, as part of all other expenses (bill of materials, manufacturing expenses, cost of delivery, etc.). The effect of such an approach on the product business entity should be dramatic, especially when the expenses charged in a certain quarter are higher than the product revenues accrued that specific quarter. The pain is sharp and is felt immediately and deeply. In that way, we prevent adding hidden (warranty) burden to the already heavy load of CS expenses. Each unit is measured accurately on what affects it the most.

How to Implement the Change³

As with every major change in an organization, there is no alternative to the managerial leadership. The kind of change needed here will require managerial ownership, as the move is clearly a top-down process. An effort to create a bottom-up process, lead by an ambitious well-wisher, rarely stands a chance of success, as the change may face a hostile reaction.

Key Decisions

Of the multitude of options presented previously, the most important one has to do with the future vision of the CS organization by the company's management. Will CS be an in-house operation or will it be outsourced? Maybe management should combine the elements of both, creating a unique combination, better suited to the particular environment.

- If we keep CS internal, what type of support contract should be the preferred one?
- If CS is outsourced, then to whom and how?
- Is the current size and mix of support personnel suitable for the future structure? If it is not, what changes are needed?

The standard TOC tools to evaluate proposed solutions—*Future Reality Tree* (FRT) and *Negative Branch Reservation* (NBR)—can help quickly screen various solution scenarios. The screening process tries to assess whether the *proposed solution can actually resolve existing problems (FRT)*, without creating even *worse new problems (NBR)*. Only solutions to emerge from both such a screening with flying colors will advance to the implementation phase.

Policies and Measurements

True to the maxim, "Tell me how you measure me, and I'll tell how I'll behave," only if proper measures are adopted can we expect the desired changes to take place relatively fast. On top of the standard TOC measures of CS contribution to the organizational bottom line, through the channels of Throughput, Operating Expenses, and Inventory

(Investment), we would like to use operational measures specifically tailored to the CS organization.

The standard measures used in CS relate to the use of different elements of the tech-support system.

1. From the point of view of CS, as usually the service event starts with a call to the Response Center (or Call Center), we would like to know what the Call Avoidance Rate (CAR) is—namely, what fraction of problems were resolved without even calling CS. The better we train the customer on our equipment, the more knowledgeable the customer is, the more available are computerized databases, and the stronger the financial incentives to avoid service calls, the larger this fraction will be. It is a measure not easy to gauge, and to get it necessitates close collaboration with the internal maintenance entity. Usually it will become visible when comparing statistical data regarding similar assemblies of comparable units of equipment.
2. *Response Center Absorption Rate is the fraction of service calls that pass through the Response Center.* Although undesired phenomenon, quite often customers bypass the call center and approach the FSE directly. It happens when it is a repeat call and they have a way to directly contact the FSE who took call earlier, the customer has a good, friendly relationship with one of the FSEs and the calls him or her directly, or (particularly in a large organization with a large installed base of equipment), a support person happens to be on the premises and can be approached directly. We would like the number of calls going through the Response Center to grow, as the smaller it is, the more it hints to a system that is not

managed by its managers, but rather by the whims of its clients.

3. As this is tricky to measure quantitatively, it can be derived if the CS personnel are required to report their *direct communications with the clients*, especially those resulting in service tasks.
4. *Response Center Close Call Rate (CCR)* is the fraction of calls that are remotely resolved (closed) by the Response Center, without the need for dispatching an FSE to the client. The higher it is, the more efficient the system is. Of course, availability of a robust expert system and an ongoing education program of Response Center personnel can significantly assist the increase of this rate.

A system is needed for monitoring the changes in all these four factors—both the direction of the changes (does it grow, stay stable, or diminish), as well as the size of them and their trend. It will give both the management and the CS staff a feeling of whether the change is moving in the right direction.

Summary

[Figure 30-9](#) presents the proposed changes in a schematic form.

In general, the changes advocated here, when applied together, are changing the very nature of the CS function; from the traditional and simple, yet expensive and with steadily declining efficiency, service organization, to a more differentiated system as shown in [Fig. 30-9](#), which embeds various levels of expertise in its different elements, and that may involve additional parties.

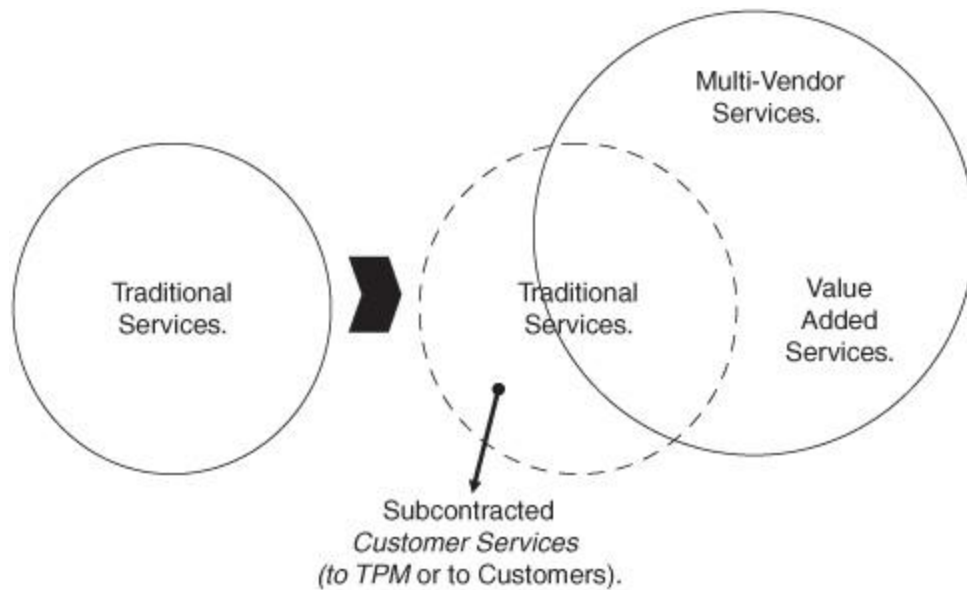


FIGURE 30-9 Shifting CS to new environment.

True, the emerging CS system looks more complex, but it has one thing going for it: It stands a much better chance of continuing to contribute positively to both the company profitability as well as to its future sales than the system it is replacing.

When the changes advocated here are seen through the lens of a system approach, the contribution of CS to the overall success of the company, which is steadily becoming a thing of the past, gets a new lease on life. I hope that it is a long and productive life.

References

Klapholz, R. and Klarman, A. 2009. *Release the Hostages: Using Goldratt's Theory of Constraints for Customer Support Management*. Great Barrington, MA: North River Press.

Turban, E. 2002. *Electronic Commerce: A Managerial Perspective*. Upper Saddle River, NJ: Prentice Hall.

About the Authors

As the President of the Goldratt Institute (Israel), **Alex Klarman**, PhD, is leading the effort to introduce TOC to and establish it as the standard management approach in Israel.

His scientific background—that of a biophysicist interested in the evolution of complex systems—as well as his industrial and educational background, including teaching appointments at the Tel-Aviv University and the State Teachers' College, in addition to years of hands-on experience in industry, makes him exceptionally qualified for this demanding undertaking.

Dr. Klarman was in charge of manufacturing in a metal industry firm for 4 years. This hands-on experience gave him a unique perspective of operations and projects—production, logistics, planning, and material management.

As the commanding officer of Dr. Eli Goldratt during decades' long service in an infantry unit of the Israeli army, Dr. Klarman became familiar with the very early concepts of OPT and TOC almost three decades ago.

In the course of the last 25 years, he took a major part in the drive to develop, disseminate, and apply TOC. His work included developing the educational materials and simulators used in various areas of TOC education, as well as the implementation work with some of the leading world-class corporations including the likes of Ford, Phillips, Intel, and Microsoft as well as many others.

Some of his works were truly pioneering efforts, like the development of TOC application in the field of intelligence analysis—already in use—or, together with Dr. Issahary of the *Dead Sea Works*, in applying Six Sigma in concert with TOC.

In the course of the last few years, Dr. Klarman has joined Richard Klapholz in the effort to develop and present a range of TOC solutions in a wide range of business activities, like sales management or customer support and service—a truly pioneering effort.

He holds a PhD in Biochemistry and Biophysics from Tel-Aviv University.

Richard Klapholz is a 17-year veteran in sales, marketing, and customer support of high technology production equipment on the global scene. He currently serves as the President of his firm's 500-employee subsidiary in Asia Pacific.

Mr. Klapholz has held various international sales, marketing, and customer support positions. All have focused on the direct distribution of strategic, high-value capital equipment, like the marketing of innovative products to the graphic arts industry through a direct sales force of 100 salespersons on a pan-European market, OEM sales to commercial and quick printers through the sales forces of Xerox (North America) and Rank-Xerox (Europe). Those sales forces included over 2000 salespersons. He was involved in sales of Dell Computers' equipment to the IT market in Israel, implementing Dell's direct sales, as well as worldwide marketing of automated inspection and imaging equipment to electronic components manufacturers through a global and direct sales channels. This involved managing a salesforce of 50 salespersons, customer support to North American manufacturers, and general management with focus on distribution to Taiwanese accounts in Taiwan and China.

He co-authored a book on sales management using TOC concepts, *The Cash Machine*. The book was published in 2004 and was later translated into Japanese, Lithuanian, and Chinese.

Mr. Klapholz is a 1992 INSEAD, Fontainebleau, MBA graduate. He became accustomed to TOC during his MBA studies and since then has become addicted to the TOC concepts and Thinking Processes. Mr. Klapholz has implemented TOC concepts in sales and customer support since 1997.

Mr. Klapholz holds a BSc degree in Electronics Engineering from Tel-Aviv University.

CHAPTER 31

Viability Vision for Health Care Systems

Gary Wadhwa

Introduction

The use of the *Theory of Constraint*¹ (TOC) in healthcare is growing rapidly; however, little has been reported in the literature. The TOC methodology has been previously applied to health care on a large scale. Knight (2003) first reported the use of *Buffer Management* (BM) in the British National Health System. Wright and King (2006) later describe the use of TOC in the British health care system in a novel form. Umble and Umble (2006) describe the implementation of TOC BM (the identification and elimination of the major causes of long waits) in three separate implementations in British hospitals. Significant improvements were achieved almost immediately using this methodology in emergency departments and the acute hospital admissions process in each implementation. For example, in the Emergency Department in Oxfordshire Horton Hospital, the pre-implementation percentage of patients processed in under 4 hours typically varied between 50 and 60 percent. In addition, the pre-implementation acute hospital admitting process regularly exceeded 4 hours and frequently exceeded the 12-hour waiting period. Post-implementation, the percentage of Emergency Department patients processed in less than 4 hours increased to about 80 percent in the next few months,

then to 91 percent, and then to 95 percent within six months of implementation. In contrast, for post-implementation results in the acute hospital admissions process, 94 percent of the patients were waiting less than 4 hours and the 12-hour waits were eliminated. Similar results were achieved in Oxfordshire Radcliffe.

The purpose of this chapter is to describe the tools, processes, and models used in an implementation of the TOC *Viable Vision* (VV)² in a for-profit healthcare practice. A VV project is an approach whereby a company—any for-profit company—maps its strategy of how to achieve net profit within four years equivalent to its annual sales today.

Our approach emphasizes TOC tools for focusing the direction of health care system improvement from top to bottom. TOC *Strategy and Tactics Trees* (S&T)³ are used to set the strategic direction from the top with tactics underpinning each level of strategic action. TOC *Thinking Processes* (TP) are used to identify core problems and the action injections needed to resolve them. Then, with constraints identified and using the *Five Focusing Steps* (5FS), buffers are put in place. Information on buffer penetration is used to sharpen the focus on specific areas for application of Lean and Six Sigma processes. The combining of TOC, Lean, and Six Sigma is a happy marriage of planning and operational methods enabling extraordinary progress. How TOC is employed and how it provides focusing guidance to Lean and Six Sigma will become clear as the story unfolds. The TOC terminology and processes have been presented in other chapters in this handbook. However, in this chapter, I briefly introduce the terminology and concepts.

The Tools for Improvement

Now we examine TOC, Lean, and Six Sigma as the main tools for strategy and process improvement in a health care practice. Each of these is a powerful capability in its own right. Combined they bring together what is needed for dramatic organization results.

Theory of Constraints

Goldratt developed a number of important TOC tools useful in system improvement. The TP are useful in identifying and resolving problems. TOC also provides a performance measurement system (*Throughput Accounting*, TA) based upon identifying and measuring a few resources (leverage points or constraints) that directly link to overall system performance. In contrast, most traditional cost accounting and some newer accounting systems (activity-based accounting) measure individual departmental performance and assume incorrectly that these reflect global performance of the endeavor. TOC also provides a number of application tools to improve the flow of goods and services and therefore Throughput of the system. New physical process improvement tools (see Sections II and III), like *Drum-Buffer-Rope*^(DBR) scheduling, *Critical Chain Project Management* (CCPM for multi-project)⁴, and *Distribution/Replenishment* for supply chain management and distribution provide system perspectives. Additionally, new marketing (Mafia Offers) and sales (Sales Funnels) approaches (See Section V) capitalize on competitive advantages. These tools have been used quite successfully in for-profit and not-for-profit organizations. One of the latest tools in TOC is called the *Viable Vision* (VV). As stated previously, in a VV implementation a company combines the use of the previous TOC tools by following logical hierarchical steps provided in an S&T to translate its current sales level into its profit in 4 years. This VV methodology

provides hope and direction for any health care system. This methodology has been applied in a small health care company, Oral & Maxillofacial Surgery, a specialty group practice with significant financial results even in the recessionary economy. The company moved from practically no profits after paying doctors and the staff to over \$3.5 million in profits per year in less than 8 years. This happened despite the time spent in learning the Lean, Six Sigma, and TOC concepts and selling them to the staff and to the doctors. Furthermore, when the focusing power of TOC points to areas where Lean and Six Sigma can be applied to increase profitability significantly, bottom-line results occur rapidly.

Lean

Lean provides a number of tools generally focusing on waste reduction across the whole value chain thus improving the flow of work (patients, in our case) through and out of the system. Several Lean and Six Sigma tools were utilized in the medical practice VV. Some of these tools with definitions are provided in [Table 31-1](#).

Lean

Lean—*A holistic and sustainable management philosophy built on minimizing resources used in organization activities and simplifying processes by eliminating non-value-adding steps with a focus on flow of parts and products from entry into the system to receipt by customers. Multi-skilled workers utilize lean methods to reduce time, blockages, and cost.*

Value stream mapping—*The process of*

diagramming and analyzing the creation, production, and delivery of a good or service through the value chain to the customer. For a service, the value stream consists of suppliers, support personnel, technology, and the provider and payment process.

Mistake proofing—Error prevention. The study of causes of errors with the focus on the elimination of the cause.

Standard workflow methods—The simplification and standardization of activities, processes, and procedures to increase workflow through an organization. A focus of these methods is the elimination of non-value-added activities.

Five S's—A set of processes (originally part of Lean) designed to clean up and make a work environment safe, efficient, and effective. These processes include: sort, simplify, scrub, standardize, and sustain.

Total preventive/productive maintenance—Worker initiated maintenance activities focused on eliminating equipment breakdowns and the continuing improvement of the equipment.

Total kit—The building of a package of items needed to support a stage in the provider-patient process. For example, having all the necessary items in place so that the doctor is able to respond to the next patient's needs without having to wait or search for any items (patient records, patient plan, supplies, instruments, assistant, etc.).

Setup time reduction—The removal of non-value-added time from the setup process.

Six Sigma

Six Sigma—A methodology to decrease process

variation and improve product quality and includes the DMAIC and DFSS methodologies.

Design-Measure-Analyze-Improve-Control

(DMAIC) process—A Six Sigma improvement methodology based on five interrelated steps: (1) design is the determination of the problem; (2) measure current performance versus the desired performance and causes of performance problems; (3) analysis to identify the core problem; (4) improve by identifying and implementing the problem solution; and (5) control by executing, monitoring, and making corrections to the new process.

Design for Six Sigma (DFSS)—A defect-prevention methodology. The process design of the value-added activities from product and process design to customer use capturing the voice of customer (VOC) and translating customer needs into quantifiable customer requirements with the objective of making processes and products robust to eliminate defects.

Quality functional deployment—A methodology to ensure that the voice of the customer (customer requirements) is clearly defined and incorporated as a customer need into the service design of the business (functional requirement). For example, a customer may not want to wait more than 10 minutes after filling out paperwork to being shown to the examining room.

TABLE 31-1 Lean and Six Sigma Tools with Brief Definitions

Lean tools,⁵ when applied to strategically important areas identified through TOC processes, cause breakthrough results. However, *the focusing power of TOC via BM and the TP are needed to focus attention where it counts.* In the current *Total Quality Management (TQM)* movement, its

tools are applied everywhere without the focus of TOC to spotlight where action will do the most good. Disappointing improvements without significant improvements in Throughput or in customer satisfaction is the result. These unfocused improvement efforts result in increases in OE only.

Six Sigma

*Six Sigma*⁶ is a statistical methodology that organizations use to reduce variations in their processes. Several health care organizations are attempting to apply these techniques combined with Lean (e.g., Virginia Mason Hospital in Seattle) but they have not used TOC. Six Sigma and Lean could benefit from the focusing power of TOC, pinpointing the best opportunities for application.

Undesirable Effects of the Current Health Care System

We now examine the areas in healthcare that need improvement. The health care system is in a crisis in this country and around the globe. To understand the problems of our current health care system better, the differing perspectives of the various stakeholders should be examined. The health care debate in the United States is surfacing an influential “voting public” perspective as people try to affect the direction of government action. However, numerous stakeholders exist. Their perspectives are important and include:

- Patients
- Doctors
- Insurers
- Hospitals

Business owners
Government

The current system pits one stakeholder against another on various vital issues.

Patients' Perspective

From the patients' perspective, the cost of health care is increasing every year; millions of people are without health care coverage because they cannot afford health insurance. Surprisingly, frequently for those who can afford insurance, the quality of service compared to other service industries is less than desirable, and the response to emergency or urgent care is poor. Patients waste a lot of time in queues waiting to get comprehensive care, being passed from one health care stakeholder to another.

Doctors' Perspective

Doctors are frustrated with the increases in their liability insurances and the low reimbursements from third-party insurers. Whenever possible, most doctors perform procedures that have low risk despite their training and experience in treating highly specialized high-risk procedures. Several towns and cities have trouble finding specialized trauma surgeons to treat patients with facial bone fractures. Sometimes patients have to wait several hours before being transported to an academic medical center for treatment. In some states a few years ago, obstetricians and gynecologists stopped delivering babies because the courts were awarding millions of dollars (exceeding their malpractice insurance coverage) for poor outcomes in medical malpractice suits. Most surgeons and physicians moved from high-risk states that award high

settlements in malpractice suits to low-risk states. Others reorganized their practices to do only low-risk procedures. Most oral and maxillofacial surgeons are highly skilled in facial injuries and reconstructive surgeries. Once they go into private practice, they soon realize both the high risk of performing these procedures and the poor reimbursement from insurance companies. As a result, they limit themselves to low-risk, high-profit, in-office procedures. While treating patients, a doctor has to weigh many variables and interactions of those variables when coming up with diagnosis and treatment plans. It takes years to learn these skills and develop intuition or judgment regarding treatment of complex diseases. The doctors are forced to multitask. They constantly face the conflict of either pleasing the insurance companies by cutting costs by not conducting expensive tests or pleasing the hospitals by ordering expensive tests. They also can refer patients with high-risk surgery to specialists versus performing the surgery themselves and facing malpractice suits if anything goes wrong. Most for-profit health care practices manipulate their mix of patients by focusing only on patients with lower treatment cost and lower risk. These choices further put them in conflict with other medical practitioners or hospitals that end up seeing these high-risk patients in emergency rooms.

Insurers' Perspective

As medical costs rise, most large insurance companies and regional HMOs are forced to focus on cost containment. They make it hard for doctors to get approval for diagnostic tests like MRIs, PET scans, and even routine CT scans—in some cases, physicians personally have to call another physician at the insurance company to get approval for diagnostic tests. Medicare and Medicaid reimbursements

are decreasing for health care providers due to budgetary cuts and financial crises in the state and national governments. Insurance companies are following suit with government cuts and further reducing reimbursements. Health care services that require more cognitive ability like family medicine, internal medicine, and pediatricians are hit the worst by these cost-cutting initiatives. These services perform very few invasive procedures. Invasive procedures are reimbursed at a higher rate than noninvasive cognitive decision-making procedures. With each reimbursement cut, these professionals are forced to see a larger number of patients in a short time to compensate for the lower reimbursement. Government is making insurance companies a scapegoat for most of the health care system problems.

Hospitals' Perspective

These cost-cutting initiatives are also hurting hospitals, especially the community-based hospitals and some teaching hospitals. Many hospitals are restructuring to remain viable. A few years ago, the hospitals were buying out private practices and developing integrated health care models. Now several hospitals are outsourcing their emergency room departments by allowing physicians to buy out the practice. Similarly, laboratory and radiology services have been separated from the hospital. Some teaching hospitals have restructured specialty services like orthopedics, neurosurgery, plastic surgery, otolaryngology and head, neck surgery, cardiology, hematology and oncology, oral and maxillofacial surgery, dentistry, and pathology, allowing the departments to run independently like for-profit private practices.

Business Owners' Perspective

Many business owners pay a portion or all of their employees' insurance premiums. As the cost of providing health care benefits increases, they must raise the prices of their products or services and thus lose their competitive advantage. Some small businesses have even gone bankrupt.

Governments' Perspective

All levels of government are under a lot of pressure from different stakeholders. Many people feel that it is government's responsibility to provide health care coverage for all the uninsured patients or to reduce the financial burden of health benefits on business owners. The business owners' assumption is that this government action will allow business owners to reduce their costs and prices to make U.S. business more competitive in the world economy.

The overall impact of the different stakeholders each making decisions for their own best interest is that the current health care system is fragmented with no one accountable to provide integrated total care for the patients. To improve care, Lean, Six Sigma, and the Balance Score Card methodologies have been applied to healthcare but they have not created breakthrough results for the overall health care system. Lean, Six Sigma, and *Business Process Re-engineering* (BPR) have resulted in process improvements. However, these local improvements have not translated into significant reductions in cost or improvement in the health area. The cost of care is continually rising and overall stakeholder satisfaction with the health care services is very low.

Defining the Goal of the Health Care System

A clear goal and a vision of the future are prerequisites for any system to embark upon the *Process of Ongoing Improvement* (POOGI). One also needs a view of the system itself. In healthcare, the goal is to increase the percentage stock of the healthy population. Several factors⁷ are necessary to achieve this goal. The key factors are preventive care and velocity of cure rate as indicated in the system model provided in [Fig. 31-1](#). The model is based upon a system dynamic model with rates of flow. The rate of identifying, treating, and preventing diseases will have a significant impact on the stock of health of our population. All strategies can be developed and stakeholder's interests can be aligned with the goal of the system.

Figures 31-1 and 31-2 give us a summary view of the medical practice system.

The model of the health care system in [Fig. 31-1](#) points out that the improvements in healthcare must be done at the system level. [Figure 31-1](#) shows that inputs affecting the disease rate (moving people from the healthy population state to the diseased population state) include the amount of preventative care, the environment of the populations, the genetics of the individuals, the lifestyle and psychosocial behavior of the individuals. The health care system provides inputs such as financial capability, system capability for quality/reliable care, system capability for rapid response to patient needs, and access to care to patients measured by the cure rate (defined as moving patients from the diseased population to the healthy population). The lower the cure rate, the higher the death rate. The model shows the significance of system capability to respond rapidly to patient needs and to provide quality and reliability of the disease management system in improving the cure rate. The goal of the health system as shown in [Fig. 31-2](#) is to transform a patient from a diseased state to a healthy state as fast as possible. The system must

be integrated with all of its parts functioning together and well to provide high quality, reliable treatment without unnecessary delays and to exceed patient expectations. The diagram further shows that all systems are made up of processes and subprocesses that can be broken down to the task or subtask levels of each stakeholder. The system must never be bogged down in the details of individual processes or lose sight of the main goal; that is, to serve the patient efficiently. The goal of aligned value chains is to satisfy patients' wants and needs

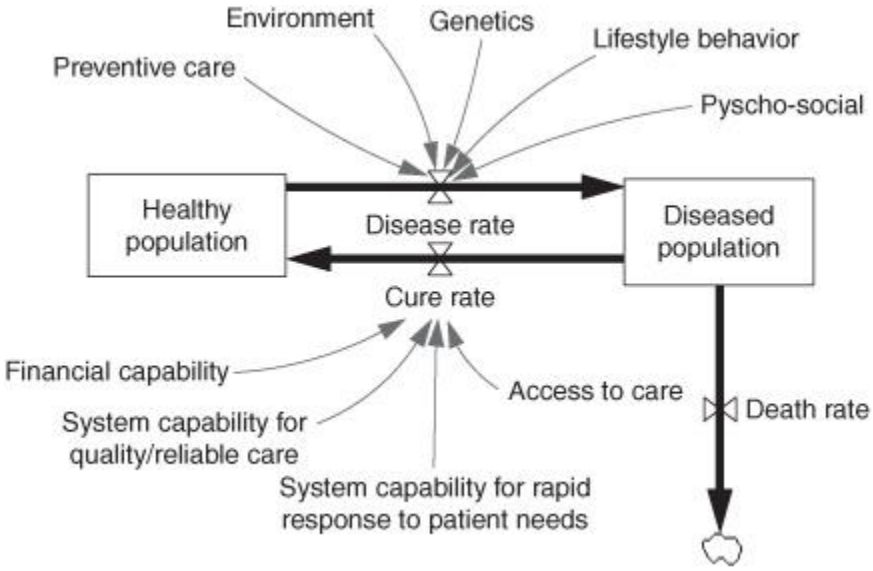


FIGURE 31-1 Model of a health care system.

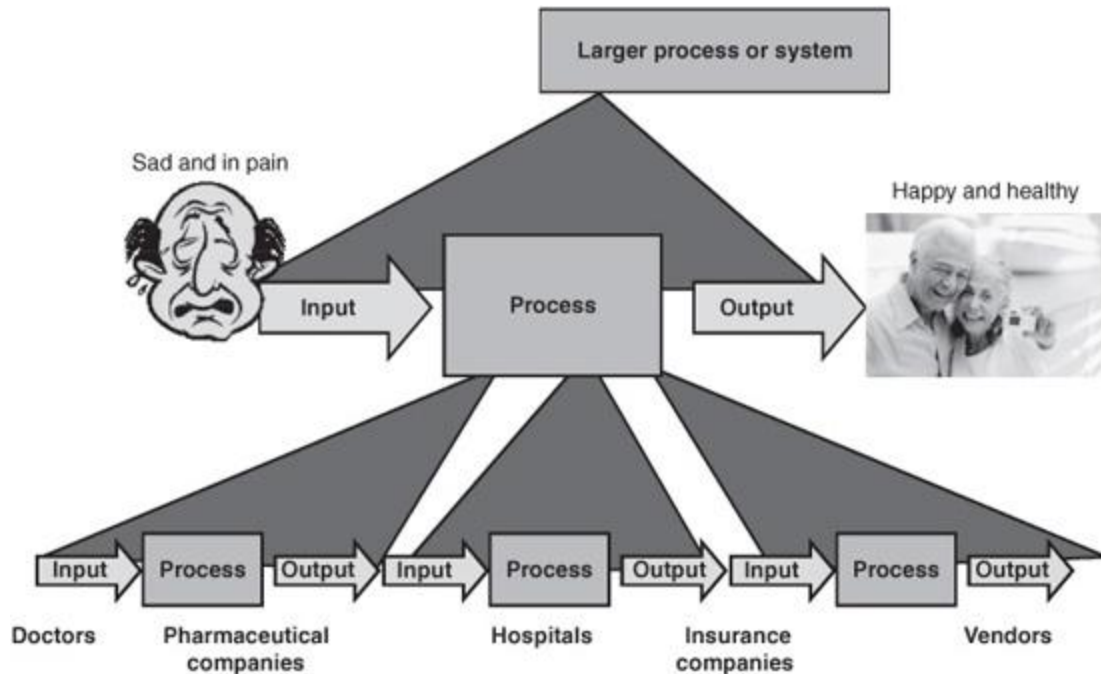


FIGURE 31-2 Process flow model of a health care system.

Improving Quality and Quantity of Patient Flow through Health Systems

Looking inside the systems pictured in Figs. 31-1 and 31-2, TOC's basic premise is that even the most complex systems have one key constraint or weakest link. For a POOGI, the *Five Focusing Steps*⁸ (5FS) (Goldratt, 1990b, 7) are useful in identifying and managing the constraint within the system and improving flow. The 5FS are:

1. Identify the system's constraint.
2. Decide how to exploit the system's constraint.
3. subordinate everything else to the above decision.
4. Elevate the system's constraint.

5. If in the previous steps a constraint has been broken, go back to Step 1, but do not let inertia cause a system constraint.

Elaborating on the 5FS

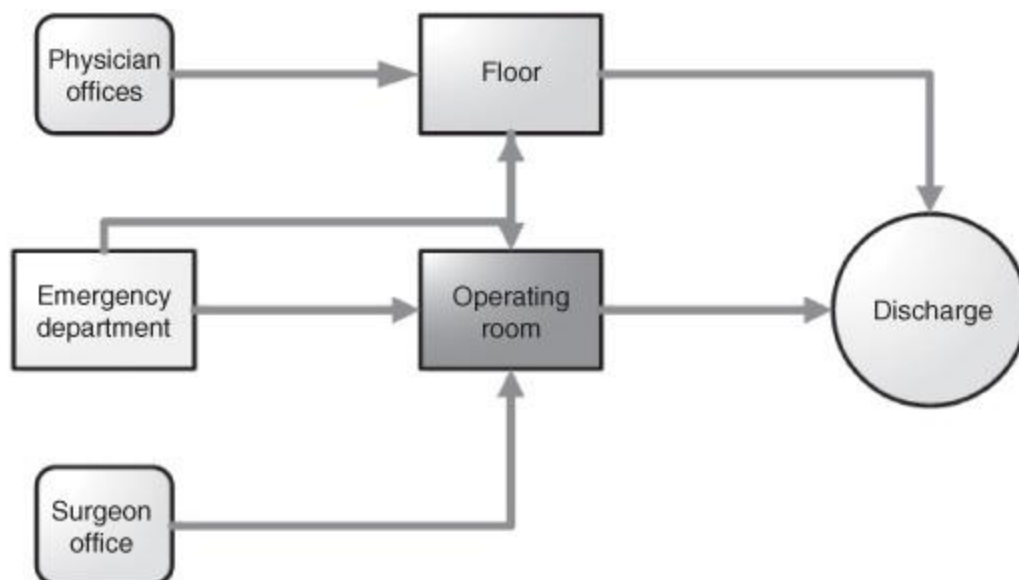
The 5FS process assumes that there is a clear goal or vision of the system's performance: the goal of making more money (Throughput) now and in the future in a for-profit business. Further, since a constraint is the weakest link in a system, it determines the Throughput of the entire system. One strategy to increasing Throughput is to be much better than your competitors are in meeting the customers' needs.

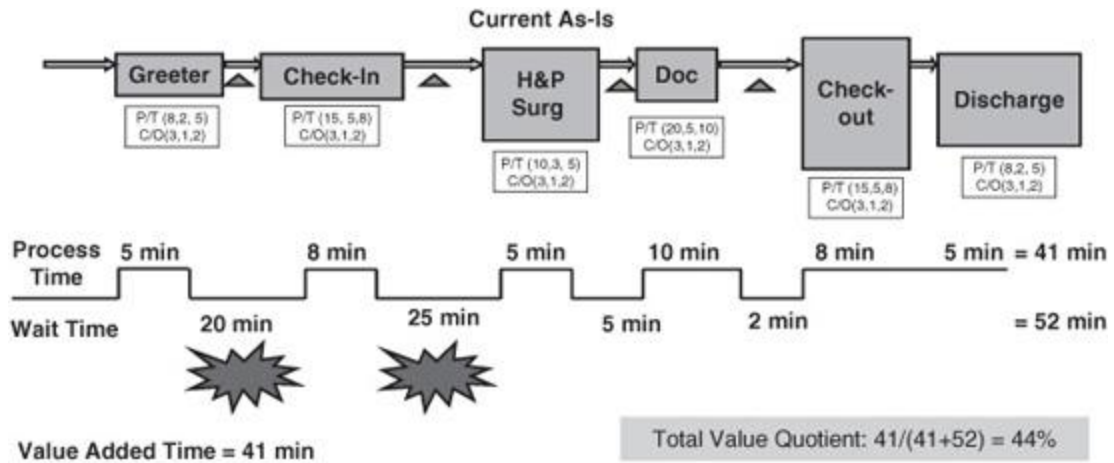
The strategy of a small or large health care system or even the entire value chain should be to develop a *decisive competitive edge* (DCE) in providing a high quality, reliable delivery and health care system. The goals of each of the subsystems have to be consistent with the overall health care system in the larger context. Example: If the specialty hospital has cardiac surgery as the financial model, it should work within the umbrella of the epidemiology or disease management model. This hospital must work to improve the velocity of flow of diseased patients through the health care system but at the same time invest in research to prevent the cardiac disease; not encourage people to get sick so that the hospital can continuously make profits. The strategies of government, insurers, hospitals, private practitioners, and businesses must also be aligned to provide high quality, reliable health care on both the prevention and curative fronts to the patients.

Step 1: Identify the System's Constraint

In a complex health care system, there is one constraint that most influences the patients' flow. It is usually the most

expensive resource: human, machine, or physical space. For example, in a small practice the constraint is usually the physician, dentist, chiropractor, or veterinary surgeon. In a larger system, it could be operating rooms, recovery rooms, emergency rooms, or CT/MRI machines. Ideally, the physician or surgeon performing the services (without whom the patients cannot flow) should be the constraint. However, due to the high investment in operating rooms, government regulations, and nursing or anesthesiology shortages, the constraint might be in one of these areas. The first challenge in the 5FS is to identify the constraint. To that end, a high-level *value stream map* (VSM) can help clarify the key obstruction or constraint to the flow of patients and information. The current or as-is VSM is provided in Figs. 31-3a and *b* and shows the flow of the patient, the value-added time (41 minutes), the wait time (52 minutes), the value-added quotient of 44 percent (the value-added time/total time in system), and the constraint (doctor) location in the system.





b. Low-level value stream map

FIGURE 31-3 High level and lower value stream maps of health care systems.

Once we understand the relationship of the constraint to Throughput and the achievement of the system's goal, we can develop protection (a buffer for the constraint), and policies and procedures for the constraint and supporting staff to maximize the system Throughput.

In [Fig. 31-4](#), the VSM can be drawn showing the complete value chain of health care businesses for a patient treatment. The flow of the patient through the supply chain will be dependent upon the dentist or lab when viewed from the larger system view. Similarly, the constraint to the flow of the patient in the hospital might be the capacity of the imaging department, the blood lab, or the recovery room nurses.

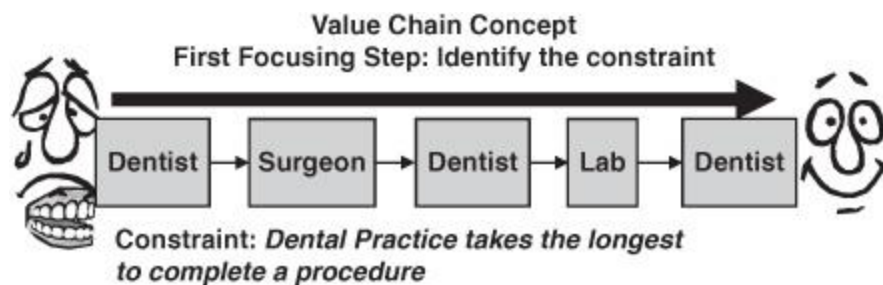


FIGURE 31-4 Value stream map of a series of businesses providing a patient treatment.

Step 2: Decide How to Exploit the System's Constraint

Once we have identified the constraint, we can determine how to make the constraint both effective and efficient. In our case, this is done by determining the best use of the doctor's time measured as Throughput/doctor time unit. Here we begin to see how combining TOC with other tools can be effective.

Example: Let us examine my practice—a dental surgeons' practice. We identify the dental surgeon as the constraint (the scarcest and most expensive resource). We next exploit the doctor's (dental surgeon's) time by using the Lean tools to effectively and efficiently use the doctor's time. We can implement the *total kit concept*, which ensures that all the documents, medical clearance, lab results, and imaging information are available to the doctor prior to seeing the patient. The Lean tools such as standard workflow and 5 S are useful in organizing our work-place to ensure everything is in its assigned place and is visually available for the doctor. Total preventive maintenance and mistake proofing ensure that the doctor's time is never wasted.

Six Sigma measures are implemented to ensure that processes are capable of achieving the desired results. DMAIC methodology is used to ensure the control of the system that protects doctor time utilization. DFSS (Design for Six Sigma) methodology can be helpful in redesigning certain processes where system capability is too low and new services have to be started to stay competitive. An example might be the use of patient focus groups to develop new services through *quality function deployment* (QFD). QFD methods can be utilized as in [Fig. 31-5](#) to identify the needs of potential patients, referring doctors,

and other healthcare stakeholders. Once the needs have been identified, the functional requirements and design parameters can be determined. See [Fig. 31-6](#). This process links new services development with a company's strategic goals.

Improving the effectiveness of the constraint is a very important part of exploitation of the constraint. The product-mix decision identifies the services that we must process through the constraint to best achieve the system goal. We have to look at the goal and the supporting strategy to determine if this is the best action to increase Throughput per constraint (doctor) time unit. In [Fig. 31-7a](#), a pie chart of the current distribution of doctor time is provided. The chart indicates the total revenues collected and approximate time for specific procedures based on time blocks in the schedule. Notice that a large amount of time is directed to facial trauma surgery (the least profitable service), while little time is devoted to wisdom tooth extraction and dental implants (the most profitable use of doctor time). In [Fig. 31-7b](#), crowns and bridgework provide Throughput of \$400 (\$1000 - \$200 variable cost ÷ 2 doctor hours = \$400) and fillings and veneer services provide Throughput of \$400 per doctor hour, while extractions/RTC provide \$350 per hour and implants provide \$250 per hour. Clearly, the surgeon should focus more on the crowns and bridgework and fillings and veneer services and less on the implants.

Data mining and an understanding of TA will help determine the services and patients that should be sought to provide consistency with the organization goal. Throughput/constraint unit time or *doctor unit time* (DU) is the key factor used in pie chart in [Fig. 31-7](#). Instead of taking each procedure and each patient that is highly variable, we aggregated the data over time. If the total collections from trauma services divided by the doctors' time utilization is significantly lower than the T/DU from

rendering services in other areas, the focus must be directed to those services with the higher Throughput.

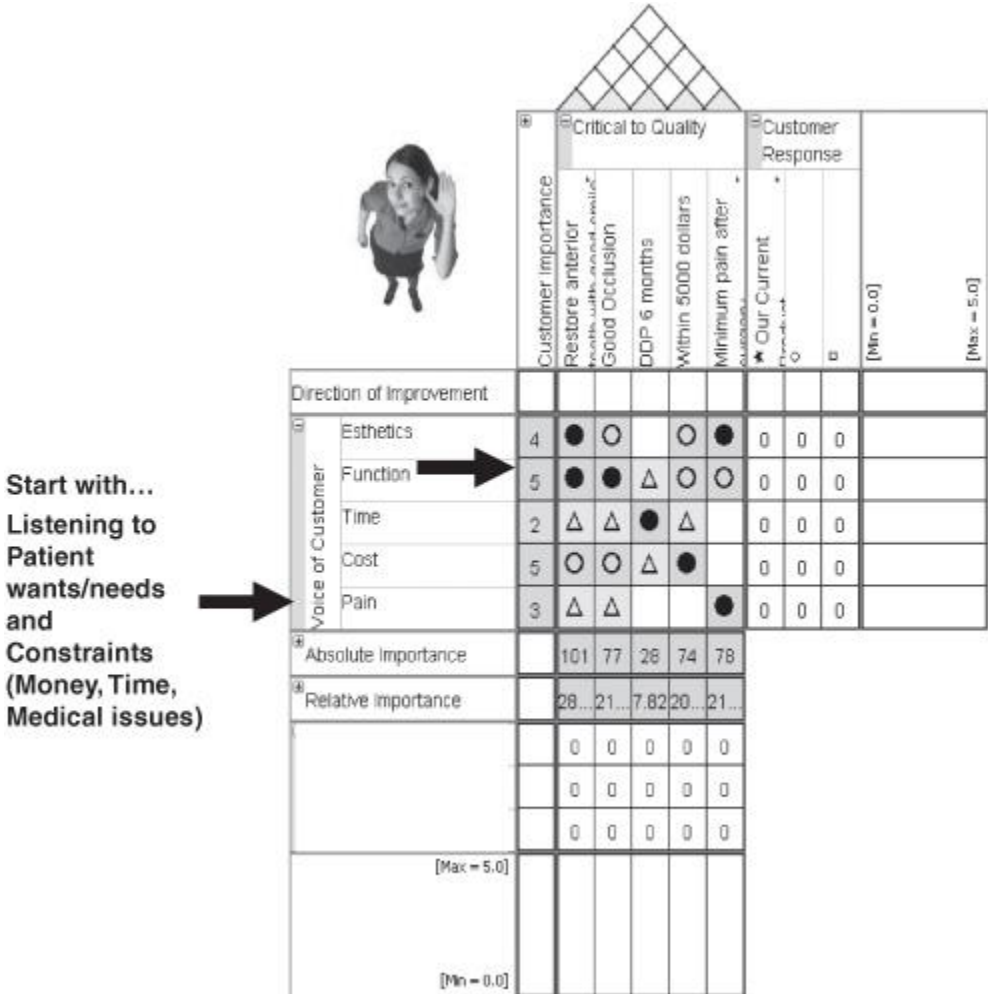


FIGURE 31-5 Quality function deployment matrix sometimes called the house of quality.

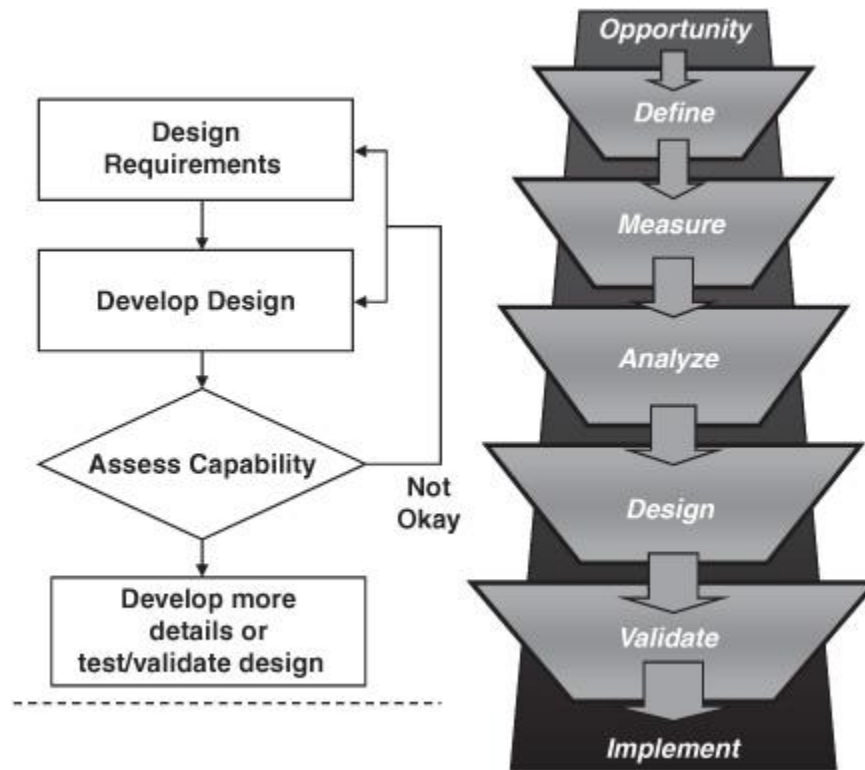
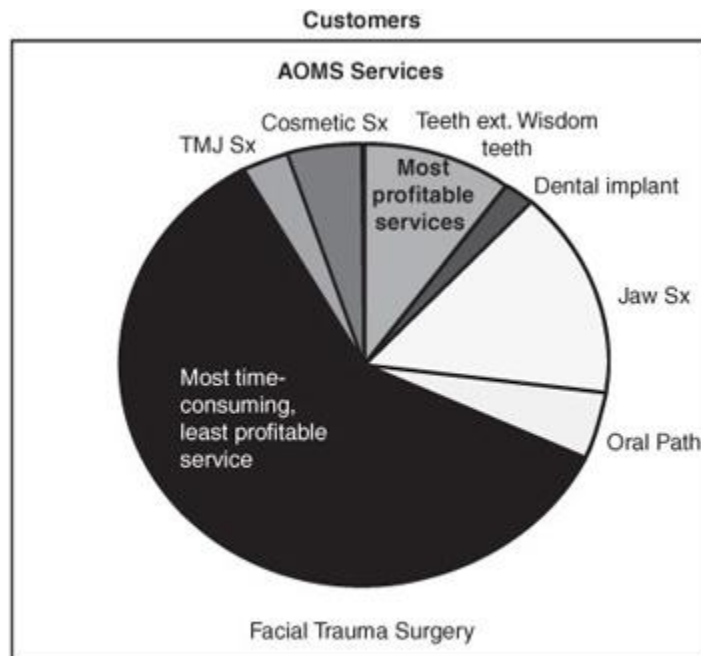


FIGURE 31-6 Design process.



a. Pie chart showing current constraint time distribution of services

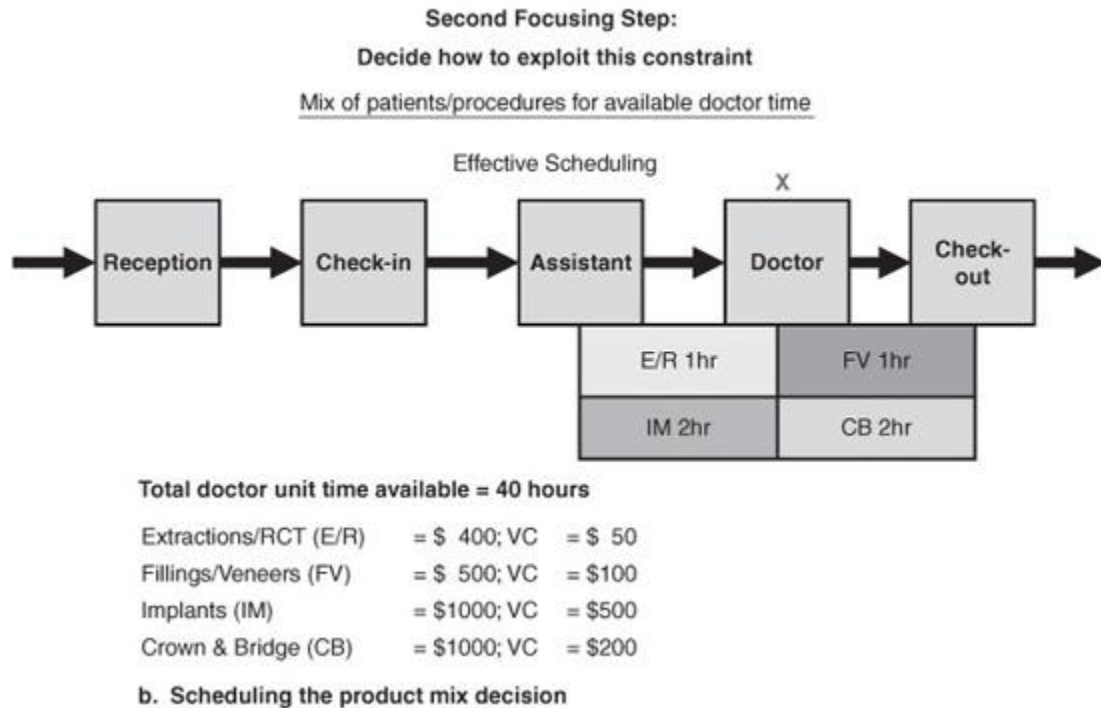


FIGURE 31-7 Current and ideal product mix.

There is another factor besides T/DU that is excessive: cost utilization of other resources. Some of the trauma cases require excessive paperwork, legal documentation, and court appearances by practice administrators or by the doctors to be paid for the services. This is a simplified version of activity-based costing, called CUT (cost utilization) in aggregate. In health care, the nursing staff, specialized billing or coding staff is an expensive resource. The cost utilization of these resources in addition to constraint resource time utilization can help with correct decision making about whether to perform or not to perform certain procedures. We also might have to make a decision about whether to perform certain procedures or refer the patient to someone else. We take into account our TA equation $NP = T - OE$. If Throughput is the function of effective and efficient use of doctor time, and OE are all the salaries, utilities, cost of inventory, etc., we must take into

account the large cost of administrative work required to do certain procedures. The increase in OE can offset the gains in Throughput.

This concept could raise many questions, but with payments capped on many procedures, for-profit health care organizations cannot survive without taking these things into account.

In the hospital context, with the operating room being the constraint, the different services such as oral and maxillofacial surgery, orthopedic surgery, neurosurgery, general surgery, plastic surgery, otolaryngology, urology, cardiothoracic, and gastroenterology should be evaluated based upon Throughput generated divided by the allocated time to the specialty services. Due to variability in patient demands in these services, most of the time these services either do not use their allocated time completely or they need additional time. Any percentage of time that was not utilized after allocation to a service or practice should be accounted as the time given to a specific service. Scheduling footprints (histories) can be developed with priority to the service that yields higher Throughput⁹ per unit of operating room time allocated and the service that has greater utilization of its block time.

Example: A community hospital has 10 operating rooms, which are the constraint in the system. The hospital is losing money and it has to improve its *net profit* (NP); otherwise, it will face closure. When we do the data analysis, we find that General Surgery has a time block of two operating rooms for two days. The Throughput or dollars collected from General Surgery is far lower than Throughput of Neurosurgery for the equivalent time block. The priority will be given to Neurosurgery. If General Surgery only utilizes 60 percent of its time and the demand for Neurosurgical patients is high, the hospital might take away General Surgery's unutilized time and give to Neurosurgery. If the hospital makes 75 to

80 percent utilization of a particular threshold, the hospital could then open the block times when another service is using less than its threshold levels. The hospital could also negotiate with staff, nursing, and administrative staff to open operating rooms for longer hours including Saturdays and Sundays. The goal should be to have flexible capacity to respond to patients' needs and wants.

In examining the data in [Fig. 31-7](#), one is cautioned to examine related factors like customer service and patient's total comprehensive needs, which must be taken into account. We cannot always look at just one procedure in isolation of total patient care. That is why it is important to take the data for each procedure and view segments of population rather than dividing the total population of patients by each procedure. It is equally important not to be guided only by the details of this analysis by looking at just Throughput per unit of doctor time because it can result in partial care and dumping of patients on other practitioners, which can have serious negative effects. Example: A practitioner selects the higher reimbursement procedures over the low reimbursement procedures and sends those procedures to other specialists. A maxillofacial surgeon in private practice can refuse to treat facial trauma patients and send them to plastic surgeons or otolaryngologists or vice versa. This might not be consistent with customer service and reputation goals.¹⁰

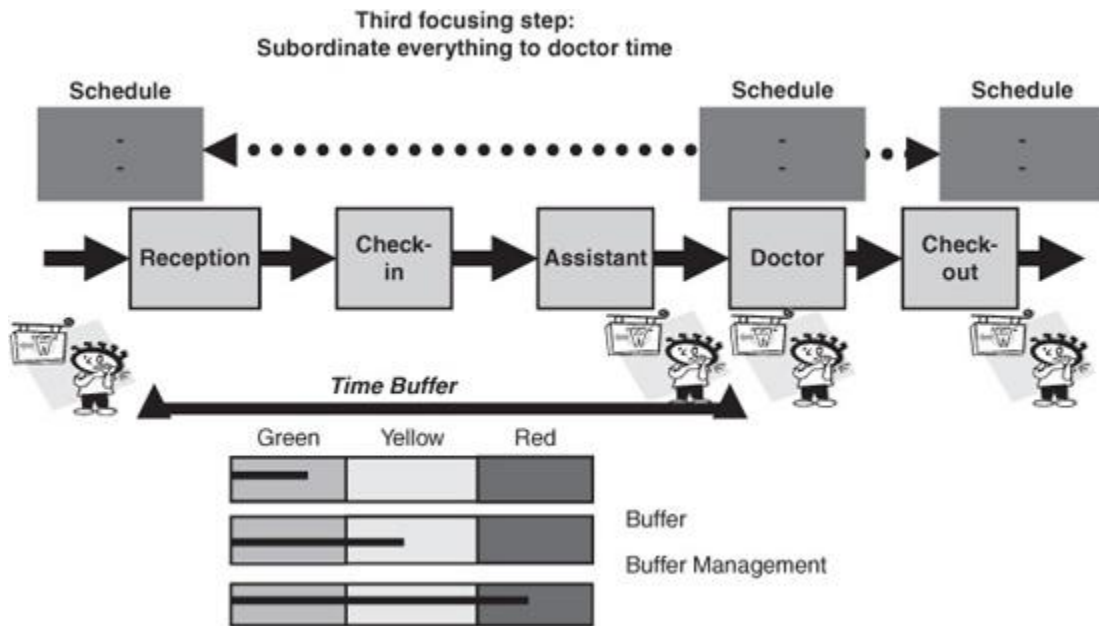


FIGURE 31-8 Scheduling the doctor’s time based on buffers and BM.

Now having some idea of the type of exploitive steps that might be taken by physicians and hospitals, we move on to examine what it means to “subordinate.”

Step 3: Subordinate Everything Else to the Above Decision

TOC offers the following methods to subordinate to the constraint: DBR, CCPM, and BM. In scheduling the patient with the doctor, the schedule procedure should be setup such that the doctor’s time is fully utilized. See [Fig. 31-8](#). Once a time in the doctor’s schedule has been identified, the patient is given an appointment (arrival) time such that he or she should arrive at the office with ample time to sign-in, show insurance card, fill out forms, be shown to the examining room, and be prepared for the doctor’s arrival. On average, the patient should have a short wait prepped in the examining room prior to the doctor’s arrival. This short wait is provided such that *Murphy may occasionally strike*, but the doctor performing his or her procedure is not

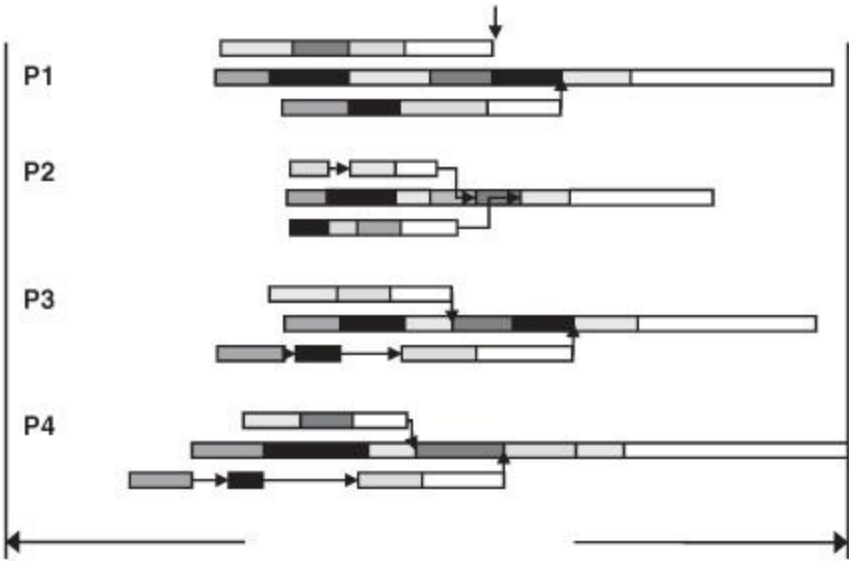
delayed. Both the appointment time schedule and the checkout schedule are derived from the doctor's schedule. All resources in the process should have ample capacity to respond to unexpected events (Murphy) and should do everything possible to keep the doctor on schedule. *This extra or protective (or sprint) capacity of all supporting resources is available in case it is needed. This is subordination to the constraint.* The buffer in [Fig. 31-8](#) is the time it takes to get the patient to reach the doctor. There is much variability in patient's arrival time, the skill sets of multiple staff members interacting with the patients, patient personalities, their medical conditions, the mental conditions of patients and staff on the particular day of interaction, etc. This interaction of variability among various factors results in delays or queuing in front of workstations. TOC provides techniques and tools of managing this interacting variability using buffers and BM reports. The buffers are placed strategically to protect the constraint resource—the doctor time. An experienced staff member takes the role of flow or buffer manager. He or she has two goals on a daily basis—to ensure that the doctor's time is efficiently utilized and the patient is not in the system longer than he or she expected. If the doctor has scheduled short procedures every 30 minutes, then we have 30 minutes to get the patient from arrival time at the receptionist to the doctor. We have buffer time of 30 minutes with 10 minutes of green zone, 10 minutes of yellow zone, and 10 minutes of red zone. If the patient arrives 15 minutes late, we must expedite this patient by doubling up the resources or doing several tasks parallel to ensure that the patient reaches the doctor in 15 minutes when he or she is done with the first procedure. *Protection of doctor time is the priority.* Similarly, the checkout or discharge is also important so that the patient is not waiting in the system after the doctor care is completed. The buffer reports tell us the trends where we have delays. If we have

check-in workstation delays, we provide staff training to identify and eliminate the delay; we then implement the Lean systems (5 S, mistake proofing, setup reduction, kitting, etc.) and re-evaluate. If we work on changing patient behavior to come on time by reminding them by phone, e-mails, text messaging, or penalizing late arrivals, then after a while, we could start reducing the buffer time when we have control over the internal variability.

In [Fig. 31-9a](#), four patients are already scheduled by CCPM. The most heavily used resource in the system is black, the strategic resource: the doctor. These networks provide the basis for scheduling the doctor's time throughout the day. If the doctor is the black resource, he or she cannot be with four patients at the same time so some shifting of the networks based on the black resource must be performed. In [Fig. 31-9b](#), the doctor time from each network is shown. Notice that the doctor is fully utilized most of the time.

In [Fig. 31-9a](#), the black resource is the doctor and he or she is supported by resources shown in other colors. The usual scenario in health systems is multitasking. The doctors and other resources are jumping back and forth to different patients without completing a single patient. This results in delays for everyone. We believe that the solution better than the DBR explained previously is CCPM. Each patient is unique and multiple providers or support staff have to work on them to get the Throughput. Multiple patients enter our system (practice) and several staff members work on these patients simultaneously. The system is prone to multi-tasking and unnecessary delays. CCPM for multiple projects with short durations can be utilized effectively to flow the patients rapidly. The Critical Constraint Resource is shown as black in [Fig. 31-9a](#). As we can see, the black resource is overlapping in all four patients. This is overly optimistic scheduling that will result in delays, and the patients will be upset. [Figure 31-9b](#) is the

first attempt to start scheduling the patient by staggering the schedule based upon black constraint resource or doctor time. Usually after three to four patients, a buffer is kept to absorb variability and Murphy which accumulated across patients. The buffers can be dynamically designed based upon customer input. If the patients start complaining within 30 minutes or 15 minutes of wait time, the psychological management of queues could be implemented. Usually for different procedures, the customers have different tolerance levels for waiting.¹¹



a. Critical chain network for four patients (P1-P4)

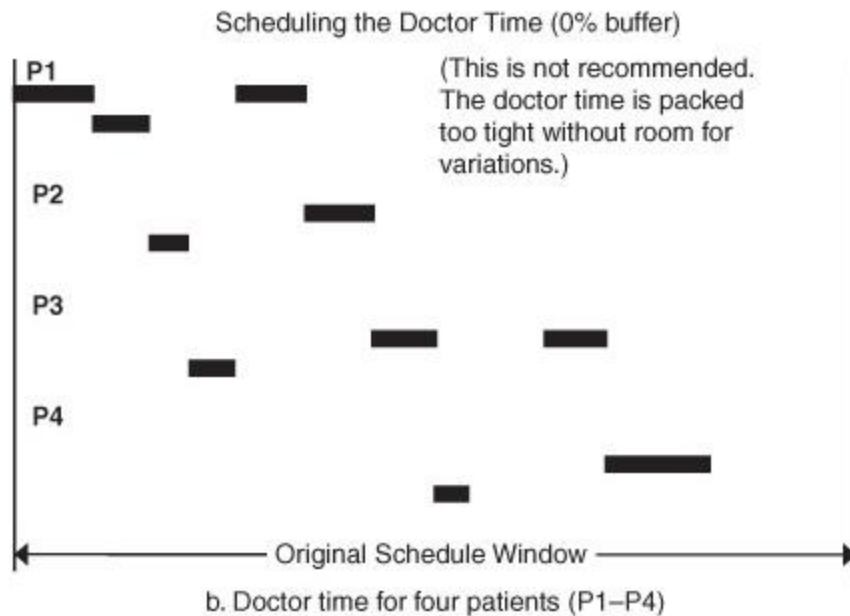


FIGURE 31-9 Scheduling the doctor’s time based on patient critical chain networks.

This mapping of the networks for all procedures can be performed manually and then the networks shifted around to fully utilize the doctor’s time, but software programs are being developed with patient care mapped out as a project. Multiple patients or patients with different needs and wants flow through our systems. We schedule a finite time that it will take to ensure that the patient exits the system within the promised time and quality of outcome. The project must start with an understanding of the patient’s expectations in addition to the medical diagnostic test results. The necessary conditions like finances (insurance, Medicare, etc.), time available and required, and patient’s existing medical condition are identified prior to starting diagnostic tests and treatment plans. After these initial steps, the best treatment designs or plans are chosen based upon the evidence-based medicine. Part of the treatment plan must take into account the patient’s inability to understand these complex concepts about their own care. Increasing the

patient's understanding or comprehension about the solutions to his or her problems will be an important step in the execution of the project for patient care so that we get full compliance from the patients.

Step 4: Elevate the System's Constraint

Elevate the constraint when we need to increase system capacity or make significant investments to offload the constraint time. To understand our investment options for elevating the constraint, we need to understand something about the TA terms to be used. We defer these now, but examine TA in more detail later in the chapter. For discussion of the Exploit step, we only need to understand the following accounting terms. Definition of these terms is woven into our discussion of the Exploit decision.

First, we must look at the impact on NP and *return on investment* (ROI) in making the elevate decision.

$$T = \text{Price} - \text{Totally Variable Cost}$$

$$\text{NP} = T - \text{OE} \quad (\text{Net Profit} = \text{Throughput}^{12} - \text{Operating Expense}^{13})$$

$$\text{ROI} = \text{NP}/I \quad (\text{Return on Investment} = \text{Net Profit}/\text{Investment})$$

Keeping this in mind, we can ensure that all of our investments in elevating the constraint will result in increases in Throughput greater than increases in OE and the ROI greater than cost of capital.

Step 5: If in the Previous Steps a Constraint Has Been Broken, Go Back to Step 1, But Do Not Let Inertia Cause a System Constraint

Sometimes the environment changes or in implementing Step 4 Elevate, the constraint moves. In these cases, one should go back to Step 1 Identify. Changes in

reimbursement or regulations from insurance companies or Medicare can cause changes in product mix, for example.

This *Five Focusing Step process* (5FS) is one of the TOC POOGIs.

Thinking Processes¹⁴ for Identifying Root Cause of Physical Constraints to the Flow of Patients

The physical constraints once identified are still difficult to manage due to conflicts in the mental models of the stakeholders. The core conflicts come between cost containment (trimming personnel until everyone is always busy, for example) versus increasing Throughput (by having protective capacity at all support functions). The other related core conflict is between local optimization (measures that focus on individual performance) versus global results (measures that focus on organization performance). These core conflicts and other conflicts are studied using the *Evaporating Cloud* (EC) technique. In [Fig. 31-10](#), the core conflict of increasing revenues versus controlling OE is portrayed as an EC with its assumptions and injections. Many of these injections¹⁵ (actions) were used in the doctor's application.

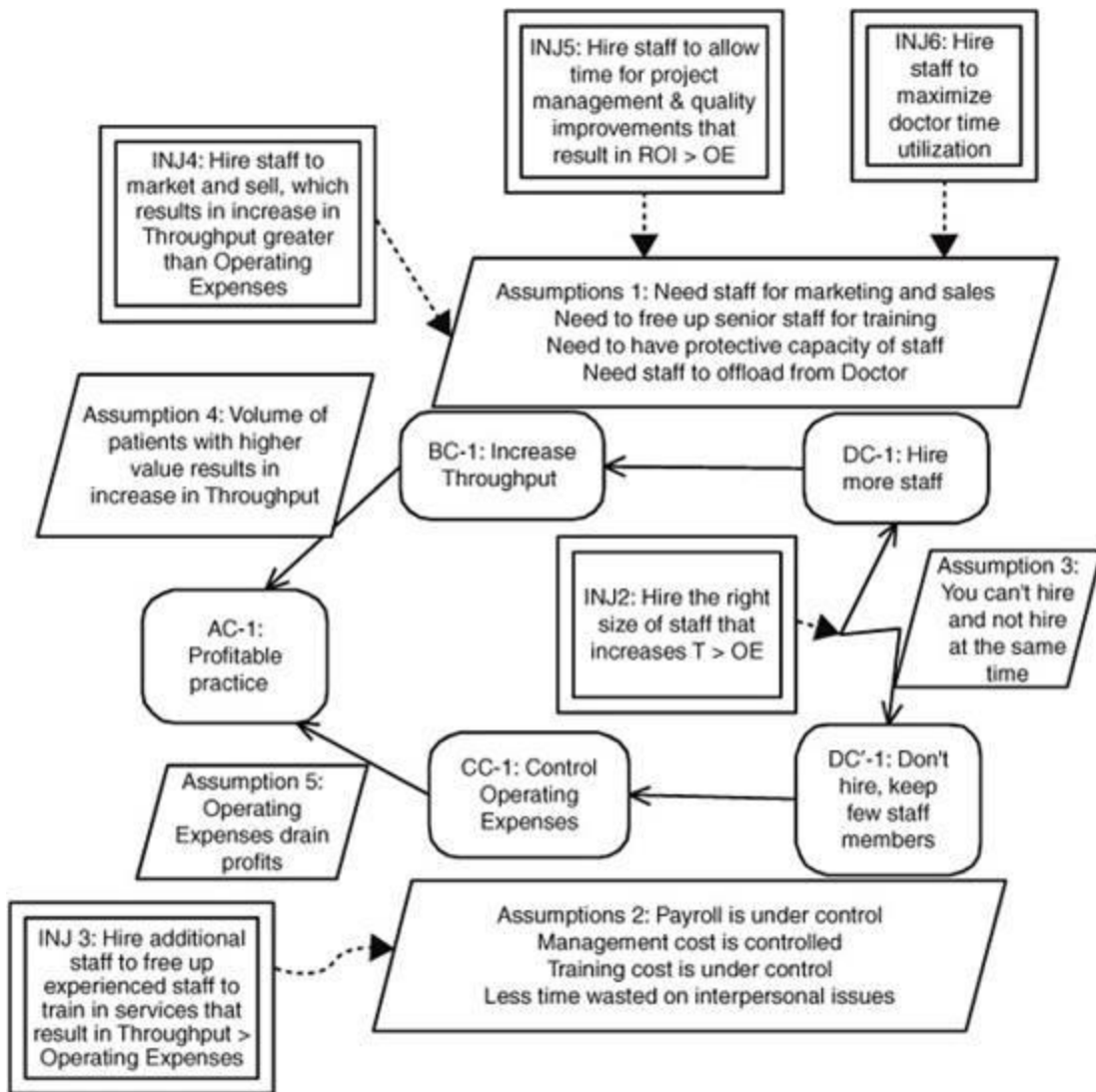


FIGURE 31-10 EC with assumptions and injections of the core conflict of hiring more staff versus keeping few staff.

Throughput Accounting for Performance Measurement and Decision Making in Health Care¹⁶

TA for health care is different from the usual TA¹⁷ in that health care service is mostly intangible. In most instances, the doctor should be treated as the constraint and the patient as the consumer. Therefore, for healthcare,

Throughput (T) means the rate of cash generation through delivery of high quality, reliable service to patients. T is the payment for services related to a specific patient minus the variable cost of laboratory work, supplies, etc. for that patient. Total T is directly related to Q (quantity of patients treated and paid for in a given time) and dollar value per patient. The quality and reliability of the process directly influence the amount of doctor's time spent managing the care of the patient.

Investment (I) is the total capital invested in designing the physical sites and delivery system of the service for the patients. It includes the cost of physical facilities, equipment, tools, IT systems, HR system, and money spent to obtain market data to develop the services for the target market. This investment is depreciated over time as OE.

Totally Variable Cost (TVC or more frequently VC) is the cost of supplies and laboratory work paid for specific tests. Since this VC varies significantly for each patient or segment of patient population, it is subtracted when the patient treatment is completed. In health care, it is not possible to focus on each patient cost due to the high degree of variability. We usually look at a segment of population served, for example, the patient population by insurance companies versus by age groups or by specific procedures.

In the hospital setting, insurance is paid by *International Classification of Diseases (ICD)* code and if Medicare population requires a lot of laboratory testing and multiple supplies, and increased length of stay, the VC will increase compared to other populations.

The symbol I ¹⁸ is used for investment that is depreciated over time, and the symbol i is used for inventory that includes medical, surgical, and office usable supplies. It can include all the unfinished treatment plans or unpaid bills from insurance companies that is similar to inventory

waiting to be worked on. OE is all the expenses to deliver services including doctors and staff salaries, benefits, leases, equipment, utilities, insurances, supplies, etc. It also includes selling and general administrative costs. I is depreciated over time as an OE. Little i is the cost of supplies, lab work, and *work-in-progress* (WIP).

As in TA, one strives to increase Throughput while decreasing Investment and Operating Expenses. The normal relationships among the variables still hold.

$$\begin{aligned} NP &= T - OE \\ ROI &= NP/I \end{aligned}$$

In health care, T can increase if we increase the velocity of understanding patient expectations, diagnose the problems accurately, create treatment plans (like design engineers), and execute the best treatment option for the patients in the shortest possible time (similar to multiple projects). The quality and reliability of patient care provided by the doctor is important in this concept because we will lose the time of our most valuable resource, the doctor, if we have to readmit the patient or we have to contact the referring doctor again and again to get the test results.

In health care, the management priorities are to:

Increase Throughput (T) > Increase in Investment (I) + Increase in Operating Expenses (OE)

In decision making and selection of patients in a for-profit health care organization, it is important to understand T/DU.

If the patient care takes too long from start to finish or a third party takes too long to pay, it will increase the OE, increase i, and decrease (T).

Staff turnover cost is viewed in terms of impact on total Throughput that is the effect on doctor time and collection of fees. To be a good decision, the hiring of staff must result in an increase in T (ΔT) that is greater than increase in OE (ΔOE) caused by the hiring of staff.

When referring procedures to other practitioners (similar to outsourcing), the decision must be made based upon overall impact on the NP of the practice at the end of the year. In decision-making, we take into account the cost of taking time off from the practice, tuition paid to develop specialized skills, investment in equipment and inventory, hiring and training of staff, opportunity cost of time allocation to providing care to select patients, marketing and sales to potential patients, and quality of service including readmitting the patients for care. The decision criterion is:

$$\text{The change in Net Profit } \uparrow \Delta \text{ NP} = \uparrow \Delta \text{ T} > \uparrow \Delta \text{ OE}$$

Decisions about developing new services in the practice must also be based upon this formula. If total T increases greater than the OE after accounting for all OE and opportunity costs, we will increase NP. Any Investment required can be expensed over time as part of OE in making the decision. If the primary care has a small laboratory to do simple tests including ECG, pulmonary function tests, blood tests, urine analysis, etc. and the increase in OE after all the investment is less than the increase in T, then the investment is a good decision. On the other hand, if the primary care wants to add imaging service in their practice, they will have to include all the I expenses in equipment, additional personnel, and time of their critical resources like the doctors learning to read CT scans and opportunity cost of not seeing regular patients while spending time in reading scans. If all this adds up to more OE than the increase in T, the decision to invest in an in-house imaging center must be abandoned.

Both *Throughput Dollar Days* (TDD) and *Inventory Dollar Days* (IDD) are valuable measures in healthcare as well. Decisions about integrating care with select specialists in health care similarly must take into account TDD. TDD is the

Throughout you would have had if a certain specialist, laboratory, or imaging center had completed their work on time. It is a penalty for lateness. IDD is made up of open treatment plans sitting queued in front of a specialist in the integrated network of health care providers with incorrect or incomplete information from the primary care provider, the laboratory, or the imaging center. It is a penalty for earliness (or doing something that should not have been done). The providers in the network who jointly treat patients can develop an informal or formal system of accountability based upon TDD or IDD.

Now that we have seen the approaches to improvement and the measurements to account for them, we will move to S&T and the approach to strategy and tactics for a medical practice.

Strategy and Tactic Tree¹⁹ to Implement and Achieve the Viable Vision

As noted earlier, the VV is an approach whereby a company maps its strategy of how to achieve NP within four years equivalent to its annual sales today. The *strategy and tactics tree* (S&T) is fundamental in mapping out a detailed strategy for achieving this outcome. In the *TOCICO Dictionary* (Sullivan et al., 2007, 43-44) the *S&T Tree* is defined as:

” A logic diagram that includes all the entities and their relationships that are necessary and sufficient to achieve an organization’s goal. The purpose of the S&T Tree is to surface and eliminate conflicts that are manifested through the misalignment of activities with organizational goals and objectives.

Usage: Organizational strategy specifies the direction of the activities that purport to address longer range problems and issues. Tactics are the specific activities needed to achieve the strategic objective involved in implementing organizational strategies. Since strategy and tactics exist and must be synchronized within various organizational levels, this logic tree translates high level strategy down to the level of day-to-day operations.” (© TOCICO 2007, used by permission, all rights reserved.)

Strategy tells us “What to achieve” and tactics tell us “How to achieve” it. Goldratt (1990a, 50-51) also points out that the most important part of any system, including a health care system (for-profit or not-for-profit), is the focus on Throughput instead of the traditional focus on cost savings. The VV is achieved by increasing the rate of flow or velocity of patient flow through the system while ensuring a high level of quality/reliability of services as measured by excellence in clinical outcomes, and total end-customer (patient) satisfaction.

The S&T tree for the VV in health care shows a hierarchical logical tree to achieve the goal. It starts with the firm agreeing on a goal. An example of a goal in a for-profit organization is to improve shareholder value. However, shareholder value can only be achieved if S1 the company is making profits over time. The profits are possible only if S2 the company is providing high value at a reasonable price to its customers. In order to develop high value services, S3 the company must develop delivery systems that provide this value and delivery systems that require highly capable people to make it happen. The highly capable people must be hired, trained, and motivated by its leadership to make it possible. Goldratt calls it strategy and tactics. The lower level specific objectives or tactics to achieve higher-level goals is the strategy.

As seen in [Fig. 31-11](#), all of the steps S1/T1 + S2/T2 + S3/T3 are necessary and sufficient—to achieve the strategy at the above level. The tree includes the strategy and tactics with the logical linkages for the parallel assumptions, necessity assumptions, and sufficiency assumptions.

Parallel Assumptions

Parallel assumptions show why tactics are necessary and how they lead to a strategy being met. At each step, we claim that the specific action plan or tactic will achieve the strategic objectives. This claim is subject to the following challenges:

1. There is no need for an action to achieve the strategy.
2. It is not possible to take the action.
3. There is another, better alternative.
4. There is a need for additional action.

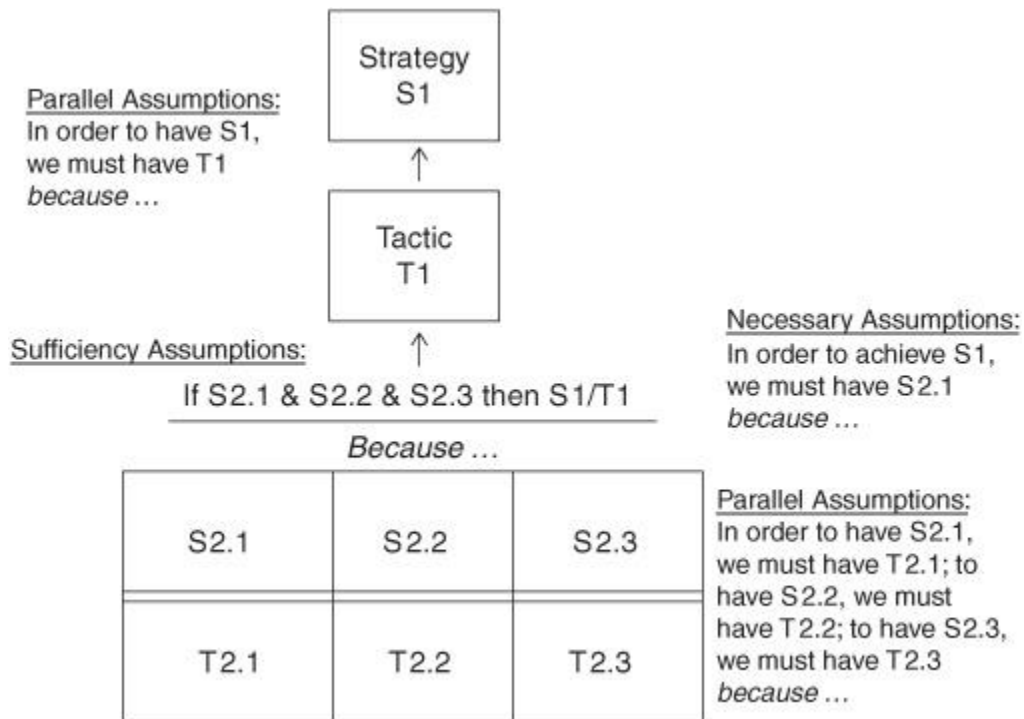


FIGURE 31-11 S&T with assumptions relationships.

How to Find Parallel Assumptions

A parallel assumption is constructed to explain the following:

1. What is currently missing that is preventing us from attaining the desired strategy?
2. Why nothing else besides what is written in tactics can achieve the strategy.
3. Disqualification of the selection of less suitable alternatives.
4. In case the tactic is challenged as a flying pig,²⁰ the lower level details substantiate the claim.

It is important to use language as a tool to verbalize these assumptions. For example: In order to achieve the strategy, I must take the action in the tactic, because The “because” response of the statement is the parallel assumption.

Necessary Assumptions

A step (for example, S1, S2, or S3) is necessary to achieve the corresponding next higher level (for example, from Level 1 to Level 2). *It is important to have an explicit explanation (the necessary assumptions) of why a given step (S1, S2 or S3 in Level 2, for example) is necessary to achieve the higher next step (Tactic x in Level 1).* There could be several necessary assumptions. It could be an answer to objections raised that this step is not necessary to achieve the next level results.

Here again the assumption should be verbalized. It should be stated as follows: In order that this step is achieved, we must do another step at the next higher level because *Again, the “because” response is the necessary assumption.*

Sufficiency Assumptions

When we claim that a group of steps (S1, S2, and S3) is sufficient to achieve the next corresponding higher-level step (S X), *we must explicitly explain (sufficiency assumptions) why all the corresponding steps of the lower-level group are sufficient to attain this step.* We write only the necessary conditions that are sufficient as a group, and an action that is necessary to achieve them. Sufficiency assumptions are expressed as: If Step 1 and Step 2 and Step 3 (S1, S2, and S3) . . . , then the higher-level step can be completed.

In order to build the tree, it is prudent to start at the higher level. Start with an objective. What is the purpose of this system? What is the reason for the system's existence?

What is the action (tactic) necessary to achieve this purpose? We write all actions necessary to achieve this purpose in the present context of knowledge. These actions cumulatively must be sufficient to achieve the objective. Verbalizing the parallel assumption, why did we choose the tactical entity to achieve the corresponding strategic objective?

An Example

Let us apply this template to an example in [Fig. 31-12](#). The S&T Tree is read from the top down and from left to right. The logic forces us to ensure that important things are not ignored or missed. Quality and reliability of service comes first, then comes marketing, and finally the growth strategy. The staff in the organization must carry out all these improvements. The processes or systems do not exist in a vacuum. Staff executes the tasks and if these tasks are tied to the system's goal, the system will succeed.

Level 1

The S&T Tree documentation contains two elements: the tree itself as shown in [Fig. 31-12](#), and an information table as shown in [Table 31-2](#). These two must be related in reading the S&T Tree. In the following example, we see at Level 1 in the strategy we have the "Health Care System Viable Vision." Relating [Table 31-2](#) to it, we have in the upper left corner a reference to Level 1 of the S&T. This reference in the upper left corner of each table links the table to the S&T Tree structure. Here, we find under "Vision" a more explicit explanation of just what the vision is. Then,

looking again at the S&T in [Fig. 31-12](#), we see the first level of tactics (2.1 Base Growth and 2.2 Enhanced Growth). Under “Assumptions behind Tactics,” a PA, we see the conditions that must be met in order for the strategy to happen: “Revenue must grow (and continue to grow) much faster than Operating Expense.” The assumptions also make clear that this must be done without exhausting the people who work within the system. Now, under “Tactic” we see the tactics that must be employed at the next level down (Level 2) of the S&T. Here we call out Tactic 2.1 “Quality/Reliable Patient Service Competitive Edge” and 2.2 calling for a special competitive edge in premium markets (Premium Competitive Edge).

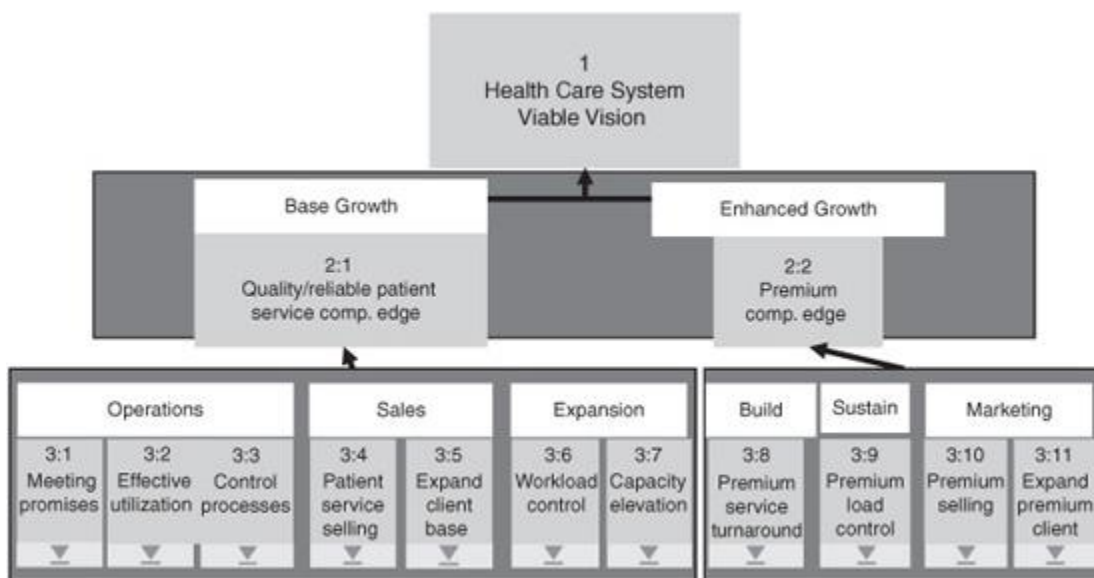


FIGURE 31-12 VV S&T, Level 1 VV, Level 2 and 3 base and enhanced growth, and Level 3.

1	Practice Vision
Vision	<p>More and more patients believe that our health care system provides the best treatment they have ever had. We have increasing capacity to service the patients with whom we enjoy working.</p> <p>As a result, today's revenues will become our net profit in 4 years or less.</p>
Assumptions Behind Tactics	<p>For our system to realize the Vision, its revenue must grow (and continue to grow) much faster than operating expense.</p> <p>Exhausting the people who work within our system or taking too high risks severely endangers the chance of reaching the Vision.</p>
Tactic	<p>Build a decisive competitive edge and the revenue generation capabilities to capitalize on it, in selected patient market segments, without exhausting our people and without taking real risks.</p>
Take Note!	<p>The way to have a decisive competitive edge is for our practice to satisfy a patient's significant need to an extent that most other practices do not, cannot, and will not do.</p>

TABLE 31-2 Strategy, Tactics, and Supporting Assumptions for Level 1

If the goal of the practice is to make money within the context of a disease management model and with full ethical responsibility, then it must increase the Throughput of the system significantly greater than the increase in OE to achieve the health care system VV in 1 ("S1"). In order to do this, the health care practice, hospital, or integrated health systems must show the capability of developing a *Decisive Competitive Edge* (DCE) over its competitors (Tactic 1, the next level down [Level 2] in the S&T). This means that they achieve breakthrough results in quality and reliability of patient care service to ensure Base Growth as in S 2.1. The system has competence to attract high premium patients for Enhanced Growth as in S 2.2.

The necessary assumption (assumption behind tactics in [Table 31-3](#)) is that quality and reliable service will improve the velocity of the flow of services without unnecessary delays and without readmissions. In addition to that, the patients will get highly reliable, quality care especially designed to capture the patients' needs and wants. This will increase customer satisfaction, which will increase the

reputation index in the market place and thus result in an increase in referrals to our system.

The sufficiency assumption (Take Note) is that it is not sufficient to have reliable services in order to make breakthrough profits. We must have additional competence to attract premium customers who pay higher than usual fees for our services. Higher premium patients result in a significant increase in Throughput without adding to the OE except for marketing and advertising expenses.

The tactical action plans to achieve S 2.1 are to implement initiatives that help develop DCE over competitors through operational excellence, sales mastery, and capacity expansion.

2:1	Quality/Reliable Patient Service Comp. Edge
Assumptions Behind Strategy	For most patients, every visit can be stressful. The more time the patient waits to complete service, the greater the stress and risk of consequences for both the patient and the practice. Therefore high quality, predictable patient service is the patient's significant need.
Strategy	A decisive competitive edge is gained by patients knowing that our health care system has a unique, systematic ability to complete all work with the fewest visits, shortest overall duration, and predictable visit times, with other parameters in place to support the best use of patient and doctor time.
Assumptions Behind Tactics	Sustaining quality and reliability in a practice is easier said than done. Having proven systems and enough support staff to back-up promises is convincing. However, in a professional practice, actual patient experience determines their willingness to return and to refer.
Tactic	Implement practice management capabilities for quality and reliability of service. At the same time, the sales and marketing procedures will help the system to grow profitability. Implement the capabilities to ensure processes stay in control, despite the demands of higher growth.
Take Note!	Building a decisive competitive edge is not easy; implementing new processes to increase doctor utilization requires a willingness to increase staff; building the capabilities to market and sell is not less difficult. However, sustaining all three elements is the real challenge.

TABLE 31-3 Strategy, Tactics, and Supporting Assumptions for Level 2

The S&T tree is a company-wide alignment, synchronization, and communication tool. The goal is to have an ever-flourishing practice that continuously and

significantly increases value for customers (patients), staff, and stakeholders. In the S&T, we agree that we are going to transform the revenues into net profits in less than four years to achieve the VV.

The tactic is to develop a DCE with recognition as a leader in providing high quality, reliable service to a select group of patients, and develop capabilities to capitalize on it without exhausting our staff capabilities or taking real risks.

The parallel assumptions or assumptions behind the tactics are that in order to realize the vision, our revenues must continue to grow much faster than the OE (hence, Throughput continually grows). However, if the growth is too fast and the staff-applied capability is not able to cope with the growth rate, the systems will collapse or oscillate and the quality and reliability of service will suffer. This will generate a negative feedback loop contrary to the vision.

Sufficiency assumptions (under “Take Note” in [Table 31-3](#)) are our reasons to believe that accomplishing a DCE will be at risk without providing another level of detail to our subordinates. The DCE is to satisfy a patient’s significant need to an extent that most other competitors cannot and will not do.

Level 2

Here we move to the next level in the tree as we see in [Figure 31-13](#). Necessary assumption or assumption behind this level of strategy is that in healthcare the patients do not visit doctors for fun; every visit is stressful and a somewhat traumatic experience. The longer the patient has to wait to complete the care (the number of visits to the doctors), the greater the stress. Therefore, high quality and reliable treatment service is the patient’s significant need.

Strategy at this level (Level 2) is to have DCE by awareness in the market that our practice has a unique,

systematic ability to complete all the necessary treatment with the fewest visits, shortest overall duration, and predictable reliable outcomes.

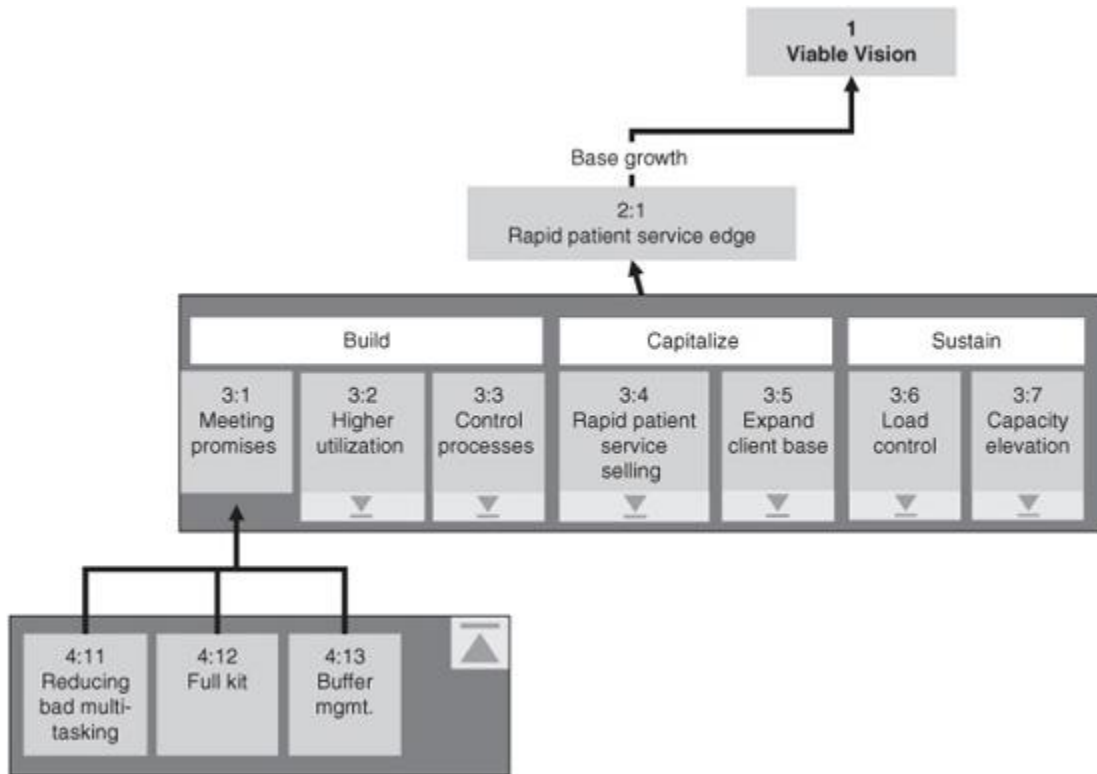


FIGURE 31-13 S&T for base growth—Level 1: VV, 2: base growth, 3: build (meeting promises), and 4: detailed tactics for meeting promises.

Parallel assumption or assumptions behind tactics will be that sustaining quality and reliability in healthcare is easier said than done. Having proven systems and enough support staff to back up our promises to patients is convincing. However, to make the systems capable of a “Wow” experience for patients is not easy. The patients do not talk about their experiences in the health care setting as much as they talk about other services. However, the future referrals depend upon word-of-mouth.

A tactic is to implement dynamic practice management capabilities for quality and reliability of services. We need to

have sales and marketing capabilities to grow the practice and to have buffer capacity to respond to emergencies and to avoid lapses in quality when rapid growth takes place.

Sufficiency assumption is that building a DCE advantage is not easy, implementing new processes requires a willingness to increase staff and increase training programs, building capabilities to market and sell is no less difficult. However, sustaining all three elements is the challenge.

This brings us to the next level on the right side of the tree.

Premium Competitive Edge

Parallel assumptions behind strategy are that to increase the probability of achieving the vision, it is helpful to have the ability to charge premiums, even on a small portion of total production. It is like hospitals providing executive health where busy executives can be treated in the shortest possible time with the highest quality for a premium. Some dental groups market to hotels, resorts, and restaurants to provide emergency care for fractured teeth, loose crowns, etc. The teeth-in-an-hour concept was promoted in different parts of the country and is another example of speed of care for a premium. Medical tourism is where patients from the United States and other Western countries go to countries like India, Brazil, and Costa Rico to get treatment. U.S. private insurance companies encourage some people to go to India for hip or knee replacement where the cost to the insurance company is one third of the cost they have to pay to providers in United States. It is cheaper for U.S. companies, but it is a premium for the other countries. They have to develop their reputation, speed, and reliability to accommodate the growing demand for such procedures. Strategy is to know the significant needs of high premium customers and design the treatment plans and delivery

systems that result in high quality outcomes in a surprisingly short lead time. It is also important to know how to market and sell to these premium customers effectively.

A tactic is to offer and deliver services for a premium with suitable training in marketing treatment planning, selling, communication, and coordination with the network of doctors, laboratories, imaging services, and suppliers.

The necessary assumption behind this tactic is that it is possible for the lead time to be surprisingly short using TOC, Lean, and Six Sigma tools. In [Fig. 31-14](#) we see, as an example, the steps leading to this dramatic reduction in lead time. Steps in Strategy and Tactic for all of the elements for enhanced growth appear in Appendix A, Tables 3:8, 4:81, 4:82, 3:9, 4:91, 4:92, 3:10, 4:101, 4:102, 3:11, 4:111, and 4:112. It is also possible to train the front office staff and doctors in identifying the right opportunities and, despite the price sensitivity and insurance industry involvement, opportunities still exist to close deals with hefty premium treatment plans.

The sufficiency assumption is that when the patient has the pressing need and is made aware of a certain health care facility capable of fulfilling that need, a sale is likely to occur.

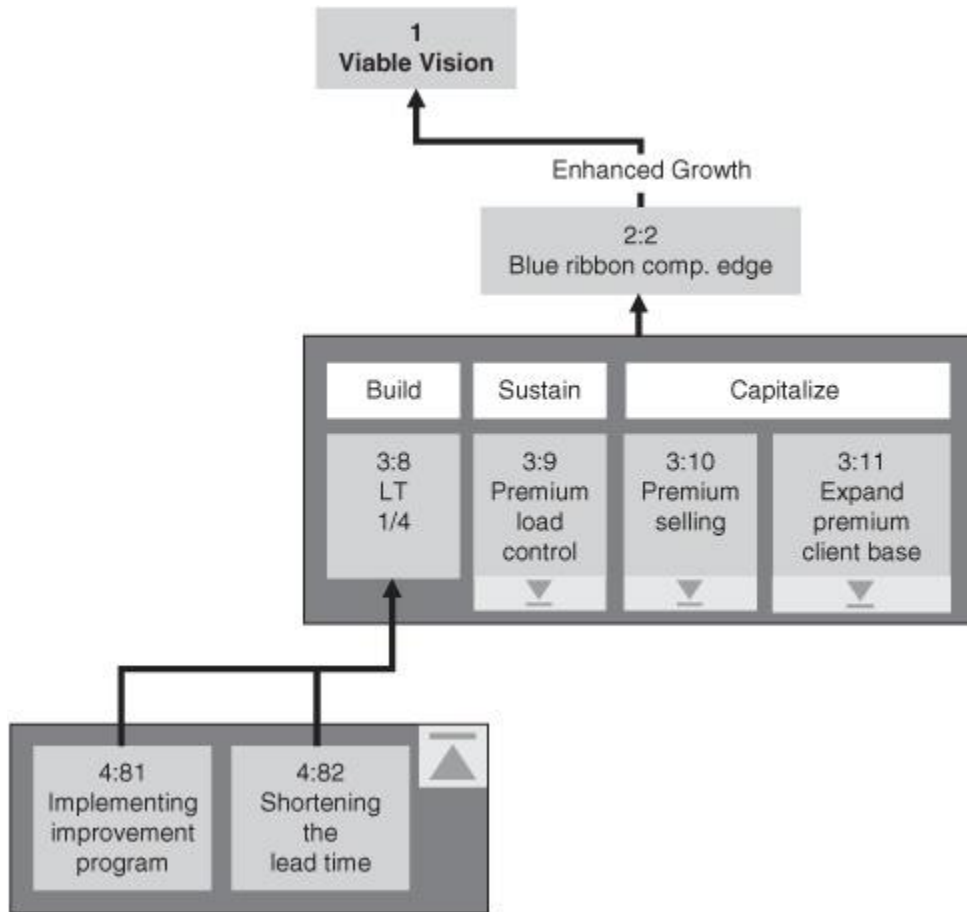


FIGURE 31-14 S&T for enhanced growth—3: build (cut LT to 1/4) and 4: implement improvement program and shorten lead time.

A Case Study of VV Success

VV was developed as a logical step-by-step procedure to help companies convert their sales figures into profits in less than 4 years. This template has been applied in some dental practices and an oral and maxillofacial surgery practice. The company grew from making minimum profits after paying the doctors to \$3.5 million in profits. All of the seven steps in 2.1 Base Growth were applied to achieve the success. The practice is again working on another vision to double the

value of the practice in the next 4 years. VV helps to develop a common language for all the staff members including doctors, medical support staff, management staff, and front-line staff. It holds different people accountable for the results.

This company set and made a goal of achieving a 100 percent increase in the value of practice in 4 years. In order to achieve this goal, the company had to attract the right kind of patients, provide high-quality care, and then flow these patients rapidly without wasting the doctors 'or the patients' time. Increasing the velocity of the flow of patients does not put any stress on doctors to perform faster. They must work to get perfection in their work so that the patients do not have to be readmitted into the system (wasting many resources). The other reason for providing high quality service is to develop good word-of-mouth referrals for future patients. The velocity of patient flow improves when we eliminate queues. The patients usually wait due to improper communication among several providers including laboratories. The company applied various tools to ensure doctor time is not wasted due to poor communications, and used TA to select the right kind of patients to flow through their systems.

Once the staff learned how to provide quality service, which is a POOGI, the staff was trained in marketing and sales. The practice made an *unrefusable offer* (URO) to the referring doctors by accepting their patients whenever the patients 'need for specialty service was identified. The practice developed the capacity to respond to their urgent patients' needs and developed a concept of providing same-day service. Similarly, the dental practices developed a system of taking patients from hotels where tourists or conference attendees stay. These patients sometimes have an emergent need for a dentist and they are willing to pay a premium fee to get the treatment done immediately.

General Discussion

The health industry has to work at a systemic level and give up its focus on local efficiencies. Global metrics like Throughput, OE, Investment, due date, and on-time performance can align multiple physicians, hospitals, test facilities, etc. in value chains with the overall objective of satisfying the end customer—the patient. TOC provides excellent tools to understand these complex systems. The Lean and Six Sigma tools are tactics that help achieve the goals of the health care systems. The TOC tools provide the focus and measurements. The focus should be on the development of human knowledge and capability of the health care organization by hiring and training personnel, rather than cutting jobs. This strategy will lead to more health care Throughput (well patients at lower costs) and hence meet the goals of the organization.

The process to improve health care systems, however small or large, is the same. This improvement methodology can be applied to small clinics as well as to large hospitals or national health services. Five improvement processes exist in TOC that are useful in a medical practice: the 5FS, TA, the TP, BM, and Critical Chain.

The patient is the key beneficiary and thus dictates the health care system design. The TOC health care methodology also encourages all health care providers to coordinate their services with the one goal of providing fast, reliable services to the patients. This integrated methodology improves Throughput of the patients through the health care system, creating more capacity to treat larger numbers of patients. The doctor is the primary revenue-generating person and hence should be the constraint resource. All other resources should subordinate their actions to the constraint and to the flow of patients through the process. It is important to apply the 5FS to

improve the Throughput and patient flow of the system. In hospitals, the radiology tools, operating rooms, etc. should not become a constraint if we want the doctors to maximize the Throughput. The TOC methodology recommends adequate support staff to protect the constraint resources and to absorb high levels of variation in health care systems. Lean and Six Sigma methodology help reduce variation and remove waste from the system, resulting in further improvements for the health system. The methodologies create high quality health care, jobs for workers, wealth for all stakeholders, and the whole value chain benefits.

References

Corbett, T. 1998. *Throughput Accounting*, Great Barrington, MA: North River Press.

Dennis, P. 2007. *Lean Production Simplified*. 2nd Edition, New York: Productivity Press.

Goldratt, E. M. 1984. *The Goal*, Great Barrington, MA: North River Press.

Goldratt, E. M. 1990a. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Crotonon-Hudson, NY: North River Press.

Goldratt, E. M. 1990b. *What is This Thing Called Theory of Constraints and How Should It Be Implemented?* Croton-on-Hudson, NY: North River Press.

Gygi, C., DeCarlo, N., Williams, B., and Covey, S. R. 2005. *Six Sigma for Dummies*. Hoboken, NJ: Wiley.

- Kendall, G. I. 2004. *Viable Vision*. Boca Raton, FL: J. Ross.
- Kendall, G. I. and Rollins, S. C. 2003. *Advanced Project Portfolio Management and the PMO*. Boca Raton, FL: J. Ross.
- Knight, A. 2003. "Making TOC the main way of managing the health system," presentation at *TOCICO Upgrade Conference*. Cambridge: Sept. 9.
- Mikel, H. and Schroeder, R. 2000. *Six Sigma*. New York: Doubleday.
- Sayer, N. J. and Williams, B. 2007. *Lean for Dummies*. Hoboken, NJ: Wiley.
- Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>
- Wright, J. and King, R. 2006. *We All Fall Down—Goldratt's Theory of Constraints for Healthcare Systems*. Great Barrington, MA: North River Press.
- Umble, M. and Umble E. J. 2006. "Utilizing buffer management to improve performance in a healthcare environment," *European Journal of Operational Research*. 174. 1060-1075.

About the Author

Dr. Gary Wadhwa is the President of Adirondack Oral & Maxillofacial Surgery Group in Albany and Saratoga, NY. He is a Board Certified Oral & Maxillofacial surgeon. He is also a Fellow of American Society of Dental Anesthesiology. He is a board certified Diplomat in International College of Oral

Implantology. He was trained in India and then Montefiore Hospital, Albert Einstein College of Medicine, New York.

He received his MBA from University of Tennessee, and Lean Implementation training and certification from University of Tennessee. He received his Black Belt in Six Sigma from American Society of Quality and Juran Institute, and Master Black Belt in Six Sigma from Sigma Pro Consulting Company. He had his TOC training with Dr. James Holt at Washington State University. He has recently started a consulting company, Strategic Planning and Practice Management Institute, with the primary objective to educate health care professionals in implementation of S&T using TOC, Lean, and Six Sigma.

Appendix A: Strategy and Tactic Tree for Viable Vision

The Appendix includes the detailed S&T Trees for a medical practice. In its first four panels, Appendix A repeats information included in the chapter text. This information is included here in order to bring together a complete set of S&T Trees and assumptions for medical practice.

It will be noted that the S&Ts proceed level by level tying strategy to supporting tactics, with tactics at one level becoming an element of strategy for the next lower level. The levels in the S&T structure are designated by the first number inside the S&T Tree boxes across each horizontal level. The number in the upper left corner of the text tables designates the level being discussed for the strategy at that level as shown in the S&T Tree. The tactics discussed in that text table refers to the tactics at the next lower level. This essentially ties the levels of strategy and supporting tactics together logically.

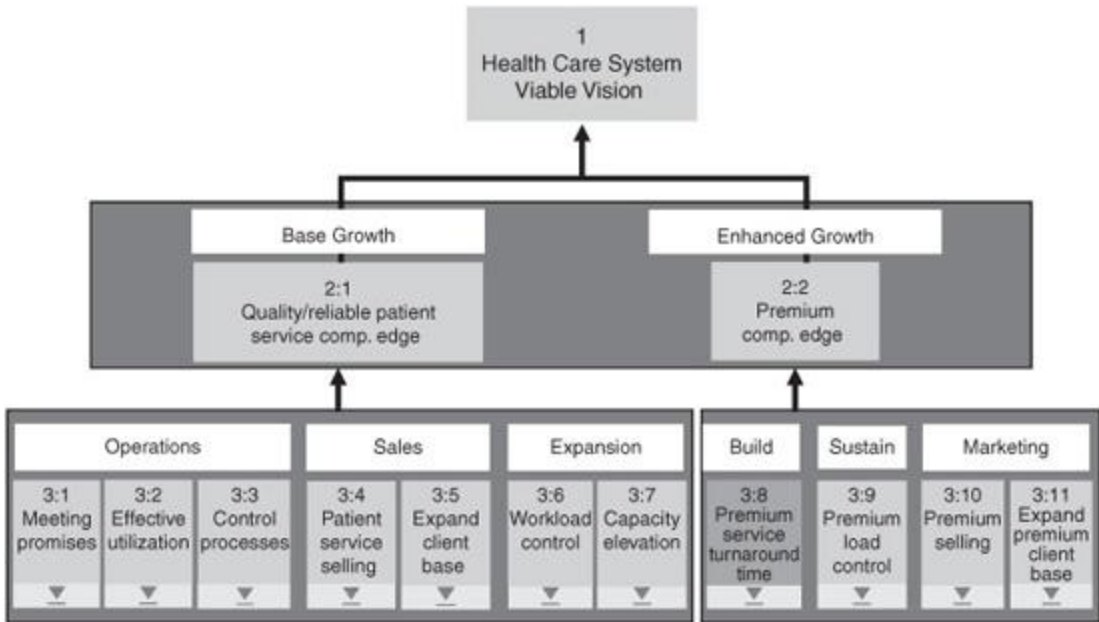
In general, the graphics that follow lay out left to right and top to bottom the S&T Trees for each element of strategic scope. Succeeding lower levels of the S&T Tree follow for

each of these broader elements of scope, showing both the strategy and tactics needed to support them. The first S&T Tree and the panel above it show the strategy in overview. The two panels immediately below show the assumptions, strategies, and tactics for each of the two major areas of direction in the strategy: 2:1 Base Growth and 2:2 Enhanced Growth.

You will notice in text boxes that the “Assumptions Behind Strategy” (Necessity Assumptions) state the reason/need for the strategy. Then under “Strategy” is the statement of what the strategy is at this level. (The strategy statement is expressed in terms of the outcomes that will be experienced after the strategy is successfully implemented. Essentially, it says “this is what things will be like when the strategy has been accomplished.”) Then, in the same panel under “Assumptions Behind Tactics” (Parallel Assumptions) are the reasons/needs for the planned tactical actions. Under “Tactics” are stated the tactical actions that are to be taken. The “Take Note” statement (Sufficiency Assumptions) in each of the panels gives cautions and advice to be considered.

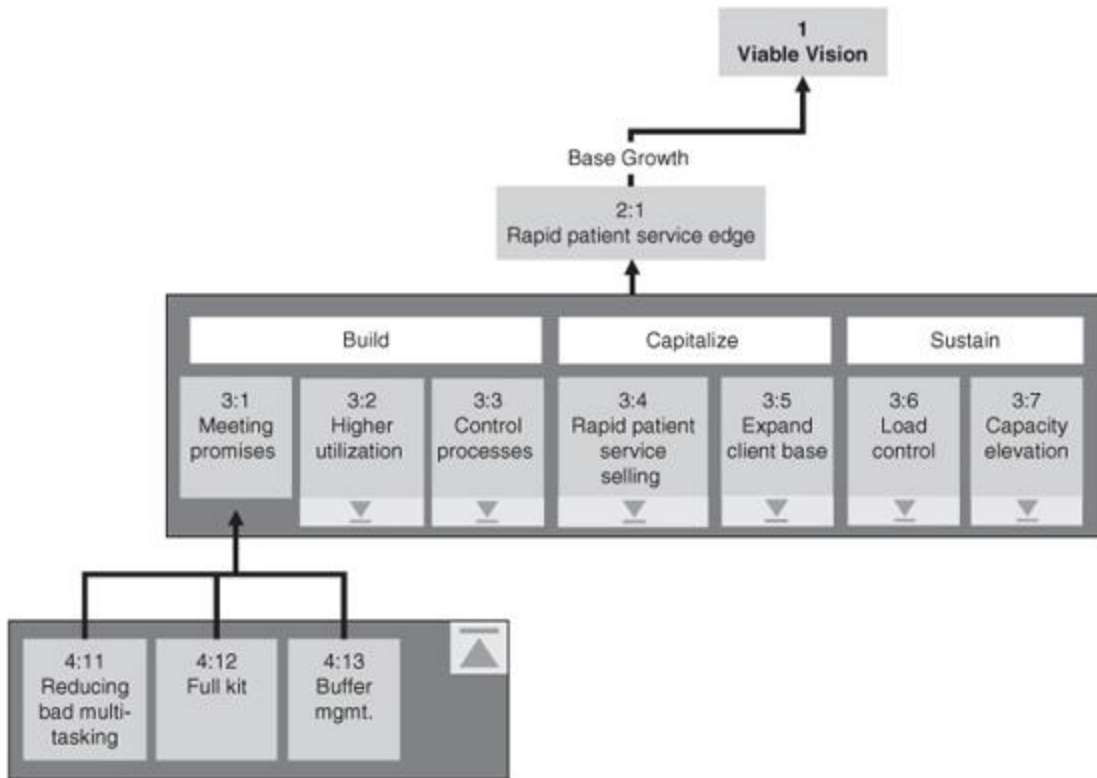
Therefore, in reading the S&T Trees that follow, you will be led somewhat by the graphics as they show a progression from left to right unveiling succeeding elements of strategic scope. Again, each of these elements of strategy is then discussed at its own level, and related to the tactics that support it one level down. The “levels” of the strategy are numbered in the S&T Tree itself as Levels 1, 2, 3, and 4. The number 2:1 indicates the first element of scope in strategy at Level 2. 2:2 indicates the second element of scope at Level 2, etc. The series begins with a panel giving the starting Practice Vision, then proceeds into the S&T Trees in this step by step, level by level sequence.²¹

1	Practice Vision
Vision	<p>More and more patients believe that our health care system provides the best treatment they have ever had. We have increasing capacity to service the patients with whom we enjoy working.</p> <p>As a result, today's total revenues will become our net profit in 4 years or less.</p>
Assumptions Behind Tactics	<p>For our system to realize the Vision, its revenue must grow (and continue to grow) much faster than operating expense.</p> <p>Exhausting the people who work within our system or taking too high risks severely endangers the chance of reaching the Vision.</p>
Tactic	Build a decisive competitive edge and the revenue generation capabilities to capitalize on it, in selected patient market segments, without exhausting our people and without taking real risks.
Take Note!	The way to have a decisive competitive edge is for our practice to satisfy a patient's significant need to an extent that most other practices do not, cannot, and will not do.



2:1	Quality/Reliable Patient Service Comp. Edge
Assumptions Behind Strategy	For most patients, every visit can be stressful. The more time the patient waits to complete service, the greater the stress and risk of consequences for both the patient and the practice. Therefore, high quality, predictable patient service is the patient's significant need.
Strategy	A decisive competitive edge is gained by patients knowing that our health care system has a unique, systematic ability to complete all work with the fewest visits, shortest overall duration, and predictable visit times, with other parameters in place to support the best use of patient and doctor time.
Assumptions Behind Tactics	Sustaining quality and reliability in a practice is easier said than done. Having proven systems and enough support staff to back up promises is convincing. However, in a professional practice, actual patient experience determines their willingness to return and to refer.
Tactic	Implement practice management capabilities for quality and reliability of service. At the same time, the sales and marketing procedures will help the system to grow profitability. Implement the capabilities to ensure processes stay in control, despite the demands of higher growth.
Take Note!	Building a decisive competitive edge is not easy; implementing new processes to increase doctor utilization requires a willingness to increase staff; building the capabilities to market and sell is not less difficult. However, sustaining all three elements is the real challenge.

2:2	Premium Competitive Edge
Assumptions Behind Strategy	<ul style="list-style-type: none"> • To increase probability of achieving the Vision, it will help the practice to have the ability to command high premiums, even on a portion of total production. • In a non-negligible percentage of cases, every practitioner involved in the accelerated delivery of services can gain. • Patients cannot get sustained economical, faster delivery (perception) of comparable service reliably from anybody except this health care provider in the community.
Strategy	A considerable portion of our service to premium patients is gained by knowing these customer needs, delivering high quality service in surprisingly short lead time, and knowing how to sell high premiums effectively to their customers.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • System can bring its lead time to be surprisingly short. • Front staff can be trained to identify the right opportunities and in spite of market price sensitivity, to close hefty premium treatment plans.
Tactic	Offer and deliver a range of short lead time services for a premium, with suitable training in sales, treatment planning, and coordination of treatment across network of doctors and technicians, including effective use of IT systems.
Take Note!	When the patient with a pressing need is made aware by the treatment planners/front staff that a particular system is able to fulfill that need, a sale is likely to occur.

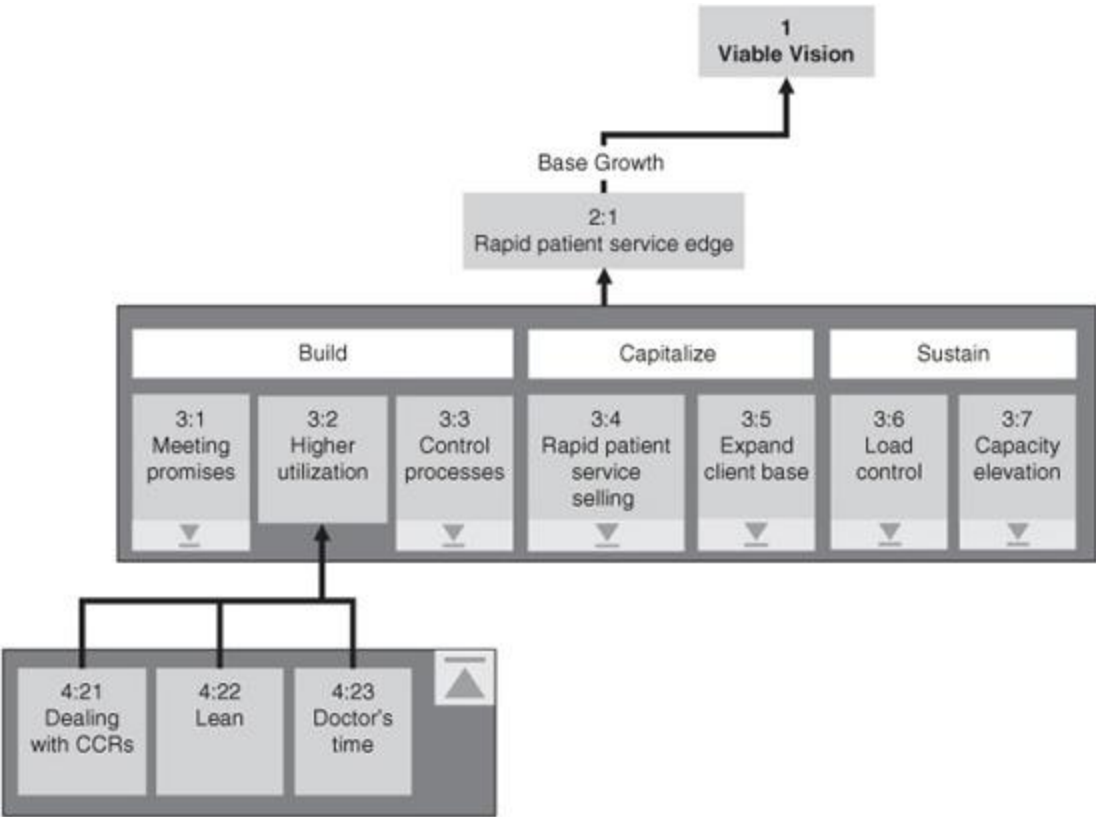


3:1	Meeting Promises
Assumptions Behind Strategy	Making a promise to patients and doing the opposite in their eyes creates an undesirable reputation. When a patient needs help, the faster the help is provided and the less of their time it takes, the happier they are.
Strategy	Patients rarely spend more than expected wait time during a visit, starting from the scheduled start time. Over 95% of patient work is completed by the original promise date, with no compromise on quality of patient care.
Assumptions Behind Tactics	<ul style="list-style-type: none"> There is a predictable level of trained staff which is necessary for rapid, effective patient flow and justifiable to the business need for profit. Systems can be easily improved to support patient flow.
Tactic	<ul style="list-style-type: none"> The practice implements rapid patient flow procedures, with the necessary level of trained support staff. Office scheduling and support system is improved to support a staff increase, by simultaneously increasing cash flow and increasing billable time.
Take Note!	To ensure an outstanding start of a major project, it is vital to ensure that each of the first substantial actions will result in immediate substantial benefits.

4:11	Reducing Bad Multitasking
Assumptions Behind Strategy	When staff are under constant pressure to work on more than one task within a short time frame, bad multitasking is unavoidable. Prolific bad multitasking significantly prolongs each patient's lead-time, reduces practice cash flow, and wastes the doctor's time.
Strategy	Patient flow is the number one consideration (the target is not to minimize cost or staff; rather, it is to complete more patient work more quickly with less stress and collect the payments much faster).
Assumptions Behind Tactics	<ul style="list-style-type: none"> • In a health care environment with bad multitasking, no one person can dedicate significant blocks of time to collections, filling schedules, etc. • Some systems are reluctant to add staff. One reason is the potential negative impact to their cash flow and profits. • More paramedical staff offloads the work from the doctor. • Experience proves that in practices with bad multitasking, giving staff dedicated time to work on cash flow and doctor billable utilization significantly increases the net profit of the practice.
Tactic	<ul style="list-style-type: none"> • The practice adds staff and dedicates people to collections and to filling the holes in doctor schedules. • Any remaining additional capacity of support staff, increased by both adding new staff and reducing bad multitasking of existing staff, is used to move patients more quickly through their visit.

4:12	Full Kit
Assumptions Behind Strategy	Current pressure often causes patients to be with a doctor without the needed preparations completed (lab work ready and quality inspected, room preparation 100% done, insurance preauthorizations, patient history checked, x-rays, referring doctor information, etc.). Sometimes this causes a patient to have to return for another visit, and this wastes the doctor's time.
Strategy	Patients are rarely seen without the necessary "full kit" preparations completed.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • The staff dealing with preparations are caught in a never ending catch-up cycle. • Hiring of staff in combination with stopping the bad multitasking frees up, for a while, ample capacity to deal with preparations.
Tactic	The practice uses the window of reduced load on the staff that does the preparations to ensure that "full kit" practice will become the norm.

4:13	Buffer Management
Assumptions Behind Strategy	The time needed for the doctor to perform individual tasks is variable. Some procedures take longer than expected, some less time than scheduled. Due to this variation, a doctor who does his or her work solely in the sequence originally planned may inadvertently cause the practice to miss its promise of visit time.
Strategy	Over 95% of patient appointment time durations are less than or equal to patient expectation of appointment duration.
Assumptions Behind Tactics	<ul style="list-style-type: none"> Doctors may prefer to work according to their own efficiency. Sometimes, meeting a doctor's efficiency conflicts with the goal of meeting patient promised on-time appointment duration. Experience in health care has proven[®] that the Buffer Management system (black, red, yellow, green), combined with short weekly meetings, leads to better on-time delivery of service to patients. <p>(See the Oxbridge case study for example. The Buffer Management system dictates the sequence of work according to the extent to which the buffer is penetrated and the extent to which expected patient time duration is in danger of being exhausted.)</p>
Tactic	<ul style="list-style-type: none"> The sequence of work by all staff and doctors affecting a patient visit is according to a single, simple priority system. Pareto analysis is documented and used to examine and fix major causes of buffer penetration into red and black zones. Multi-disciplinary meetings are held to a maximum of 1 hour per week, with actions identified and implemented before the next meeting.



3:2	Effective Utilization
Assumptions Behind Strategy	The lower the utilization of billable practitioners, the greater the danger of negative cash flow and of exhausting the resources.
Strategy	Number of billable hours is increased by at least 25%, with billings at least double any increase in expenses, without increasing the total number of service provider hours committed to the practice.
Assumptions Behind Tactics	There are many reasons why a service provider does not bill all available hours. However, there are only a few, at any point in time, that are real leverage points for significant improvement in profits.
Tactic	The system dedicates staff to focus on three major opportunities for profitably increasing utilization—addressing capacity constraints, applying Lean concepts, and improving doctor's use of their time.
Take Note!	It is impossible to achieve and sustain results without clarity. The stakeholders of the system must go out of their way to ensure that their staff fully understand the reasons for a change in staffing or procedures, the financial implications, and the expected results.

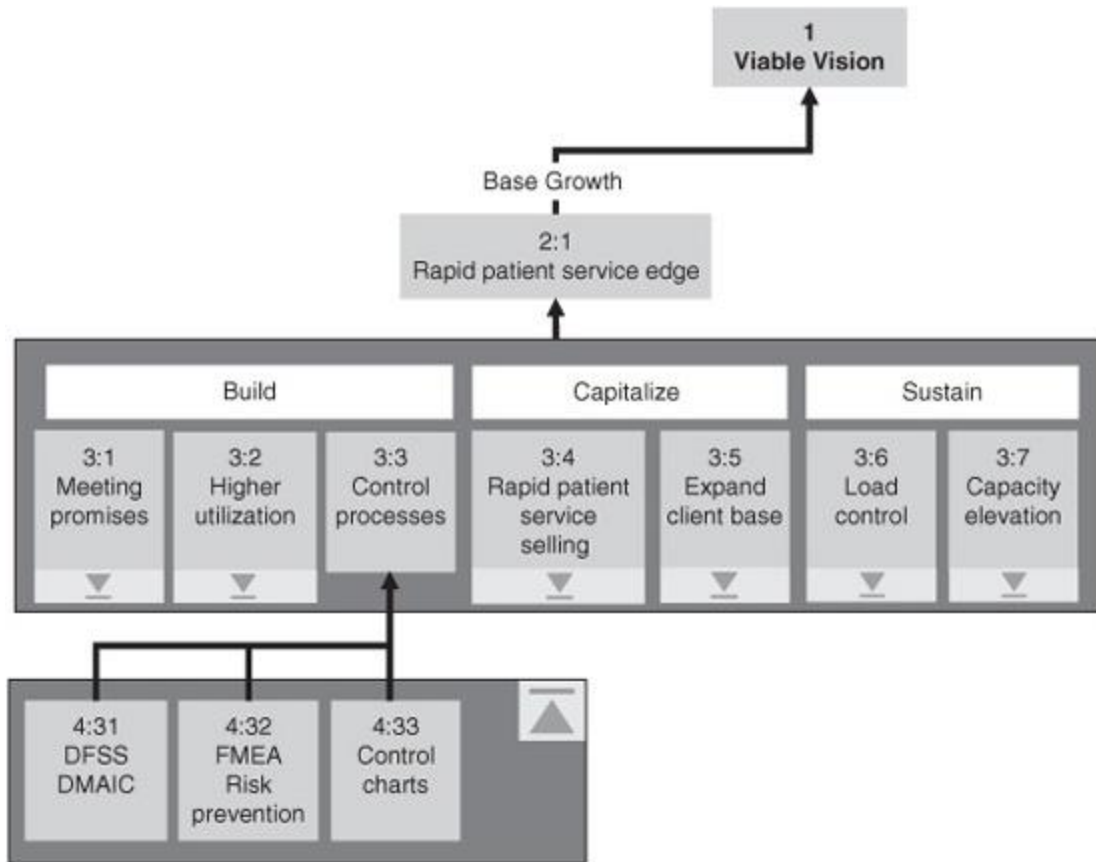
4:21	Dealing with CCRs
Assumptions Behind Strategy	There are capacity constrained resources (CCRs) that prevent achieving a much higher billable hour utilization from billable resources.
Strategy	A higher percentage of every service provider's available time is spent doing billable work.

Assumptions Behind Tactics	<ul style="list-style-type: none"> • When a doctor is unable to perform more billable work, there are two typical reasons that relate to the CCR concepts. One is that they are the CCR, with a pile of work to do that consists of billable and nonbillable activities. In the other case, the support staff is the CCR, which leaves the service provider with less effective use of his or her time. While "full kit" (see 4.12) takes care of some of these problems, other issues remain, such as poor technology, insufficient skills or empowerment of support staff, etc. • Earlier steps taken are typically sufficient to prevent the CCRs from jeopardizing on-time performance.
Tactic	<ul style="list-style-type: none"> • Support staff CCRs are identified and effectively removed. • Nonbillable service provider work is offloaded. • Empowered staff develop or hire necessary new skills and technology to deal with CCRs or help achieve higher billable hours. • Throughput accounting principals are used to ensure profitable hiring and investment decisions.

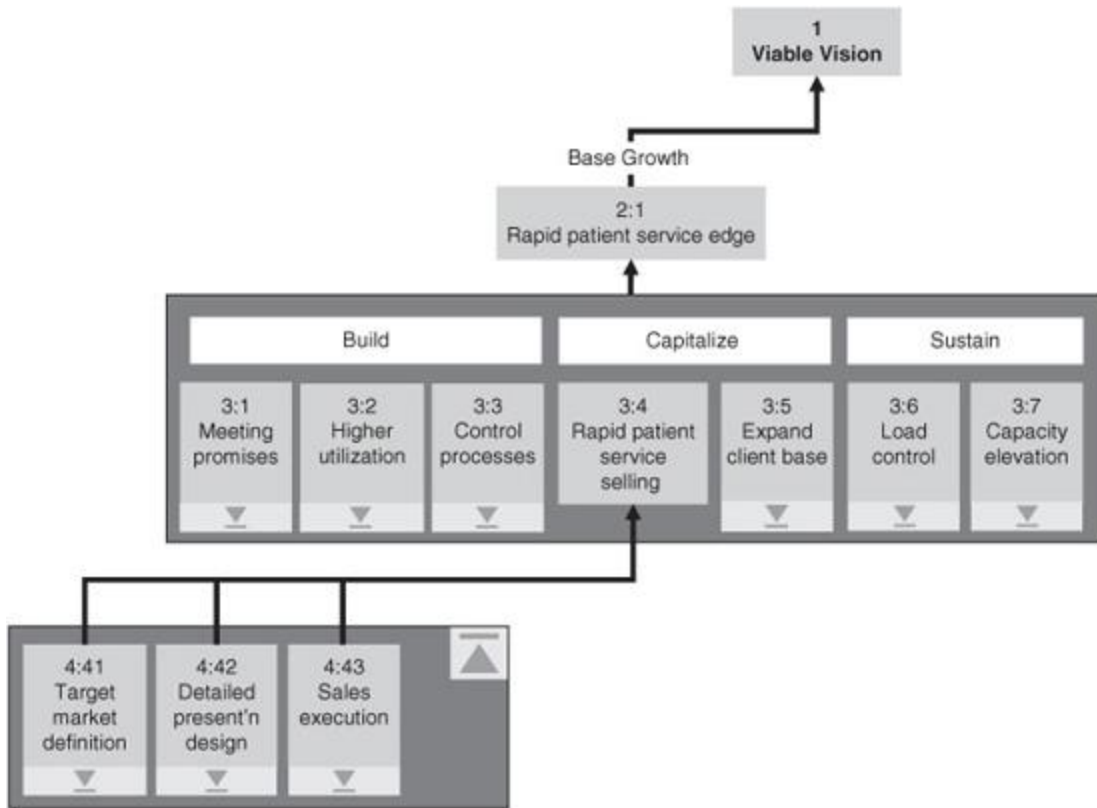
4:22	Lean
Assumptions Behind Strategy	Support staff in a practice often do not have the training, experience, or time to continually improve productivity.
Strategy	Doctor utilization and billings are significantly improved through limited application of Lean/value stream thinking.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • There are support factors (such as patient set up time, doctor wait time for staff to do activities, availability of materials in every room, availability of consultation room) that affect doctor utilization. • There are doctor work content factors (such as doctor setup time, space utilization) that affect doctor utilization.
Tactic	<ul style="list-style-type: none"> • Doctors agree to participate in and support a Lean effort to improve their billable utilization. • A "Lean Team" is identified to learn and apply Lean techniques according to Lean Practice Management principles. • The Lean Team offloads other responsibilities to allow them sufficient time to implement Lean application. • Over time, other staff in the practice are trained in, and given immediate projects to implement, further Lean techniques focused on significant profitability opportunities.

4:23	Doctor's Time
Assumptions Behind Strategy	Doctors sometimes like to, or feel compelled to, do too much nonbillable work.
Strategy	Doctors willingly trade low value nonbillable work for higher value billable work.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • The higher a doctor's number of billable hours, the more able the doctor is to choose the work he or she likes to do and the work that is more profitable for the practice. • The more profitable a practice is, the more quality time a doctor can have. • With a different approach to staffing, it is possible for doctors to do less billable work and have results that are more satisfactory.

Tactic	<p>Staff, together with the doctor, dedicate time each month to examine how to:</p> <ol style="list-style-type: none"> 1. Help trade the doctor's nonbillable time for billable time. 2. Identify opportunities and capture (sell) more of the higher premium work. <p>From each meeting, the target is to have one idea identified and implemented before the next meeting</p>
--------	---



3:3	Control Processes
Assumptions Behind Strategy	Without simple, robust processes to ensure predictable outcomes, a system is at great risk of not being able to sustain profitability and patient satisfaction.
Strategy	Rework and risk are minimized. Patient satisfaction is greater than 95%.
Assumptions Behind Tactics	The skills to sustain quality do not exist within most health care systems today. Without the skills to monitor processes constantly over time, quality will degrade to an unacceptable level.
Tactic	Train a team in the quality skills to define, monitor, and achieve predictable patient outcomes.
Take Note!	The body of knowledge for quality/Six Sigma is huge. Practice staffs are not unlimited in ability, time, or cost. Training must be focused and kept simple in order to succeed with sustainable quality.

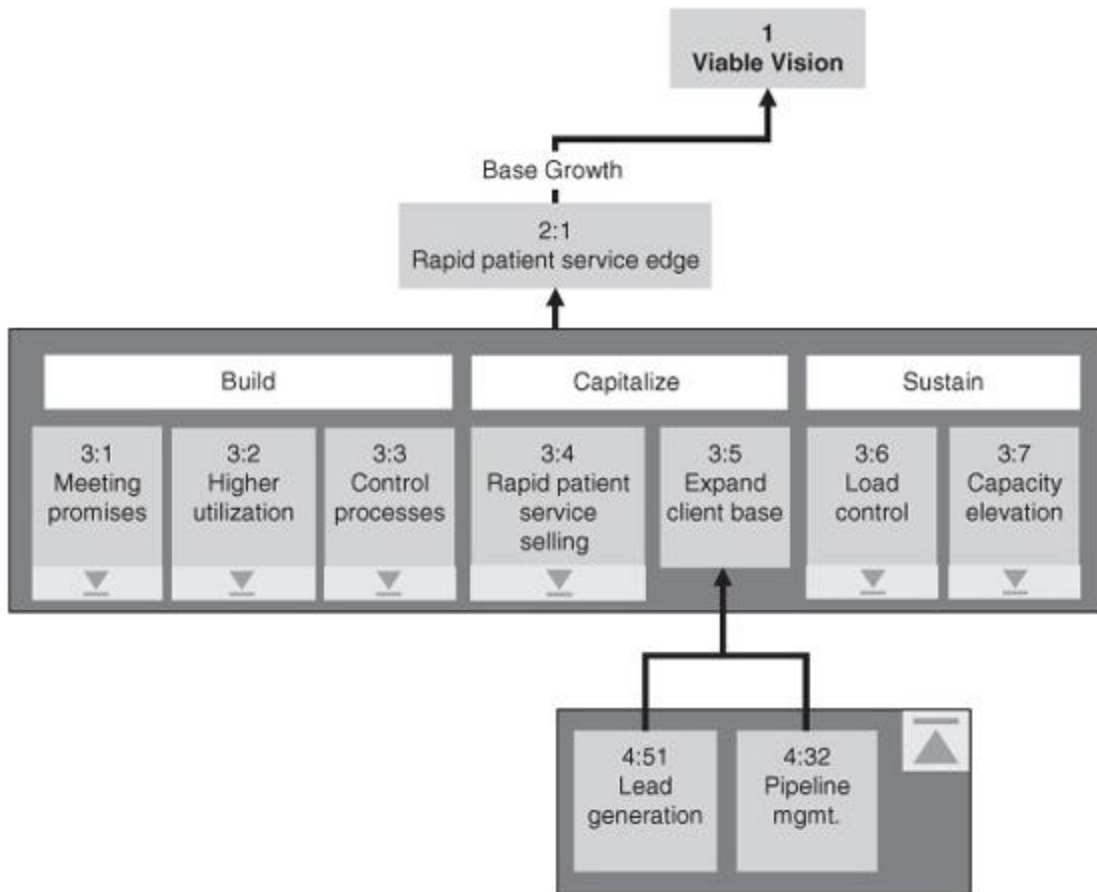


3:4	Patient Service Selling
Assumptions Behind Strategy	<ul style="list-style-type: none"> The required changes in the practice's approach to capitalize on remarkably better service (the Predictable Patient Service offer) are different in nature from the changes the practice did in the past (new procedures or new products). Leaving the positive impact of remarkably better predictability to the natural word of mouth of patients will take too long to reach a decisive competitive edge.
Strategy	Revenues generated by the Predictable Patient Service offer are increasingly growing.
Assumptions Behind Tactics	The changes in the marketing and sales approach require time and there is no time to lose. The improvements implemented in practice management reduce patient duration and expose capacity. Delaying the sales effort may erode the doctors' and staff's confidence in the solution.
Tactic	From the outset of the Vision project, the practice aligns its staff sales approach and trains staff to take full advantage of the Predictable Patient Service offer.
Take Note!	Having a competitive edge that is predictable service-based and simultaneously having the capacity to address it is a paradigm shift for support staff that are not trained in solution selling. Some staff may even feel very negatively about sales.

4:41	Target Market Definition
Assumptions Behind Strategy	Pursuing the wrong patients is not just a waste of valuable resources (money, sales capacity, time, etc.) but also can lead to the "conclusion" that the offer and its underlying solution are invalid.
Strategy	Staff and doctors agree on which patients to pursue with the Rapid Patient Service offer.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • There are patients where Rapid Patient Service is not a significant need. • There are patients where Rapid Patient Service is a significant need. However, they are too risky or require excessive efforts to work with.
Tactic	<p>Target patients are defined according to conditions that are:</p> <ul style="list-style-type: none"> • Easily checked, and • Relate to a non-negligible number of patients. <p>The conditions prioritize patients according to:</p> <ul style="list-style-type: none"> • The degree to which rapid patient service is a significant need; • Referral source; • The estimate of the ratio efforts/returns; and • The degree of business risks.

4:42	Detailed Presentation Design
Assumptions Behind Strategy	<p>When the details of an offer are not clear, staff will not sell it, because they think the risk is too high, the benefits are trivial, or they simply do not understand it.</p> <p>When the details of the offer presentation are not compelling to the patient, the patient may not commit to the work.</p>
Strategy	The practice has a Rapid Patient Service presentation that results in the patient booking the work within hours or a week of the first appointment almost all the time.
Assumptions Behind Tactics	<p>To construct a good offer presentation, four elements must be thoroughly understood:</p> <ul style="list-style-type: none"> • The net benefit for patients relative to traditional practices. • The benefits to the practice. • The risk for the patient (relative to traditional practices). • The risk to the practice if they do not meet the expectation. <p>Ensuring the benefits provides the detailed backbone of the offer. Mitigating the above risks provides important details of the offer.</p>
Tactic	<p>A team is empowered to construct the details of the Rapid Patient Service presentation, maximizing the benefits (to both the patient and the practice) and minimizing the risks.</p> <p>The team creates a detailed presentation that describes the risks of delays in completing the necessary work, the potential damage to patients when their appointment time is not respected, the systems the practice has put in place to ensure success, and a realistic expectation based on actual performance.</p>

4:43	Sales Execution
Necessary assumption Behind Strategy	Conventional sales methods are not effective enough to capitalize on a competitive edge that stems from anything other than providing the patient services alone. Not all staff know how to or are capable of effectively selling.
Strategy	Increasing numbers of new patients are buying in to the Rapid Patient Service Offer.
Assumptions Behind Tactics	<ul style="list-style-type: none"> It is possible to switch some customer service staff from the conventional mode of sending gifts, reminder cards, etc. to the very different mode of selling a service concept that treats the patient's time and need to complete service quickly as significant needs. Acquaintance with the patient's decision process together with the experience of selling a decisive competitive edge offer can be used by customer service/sales staff to document and follow a simple, successful sales process.
Tactic	<ul style="list-style-type: none"> Define the sales process—what the staff should do, at which stage, how (using standard tools), with whom and by whom in order to bring an identified patient from "ignorance" to closing a deal. Train, coach, and handhold the customer service/sales staff in selling the Rapid Patient Service offer.

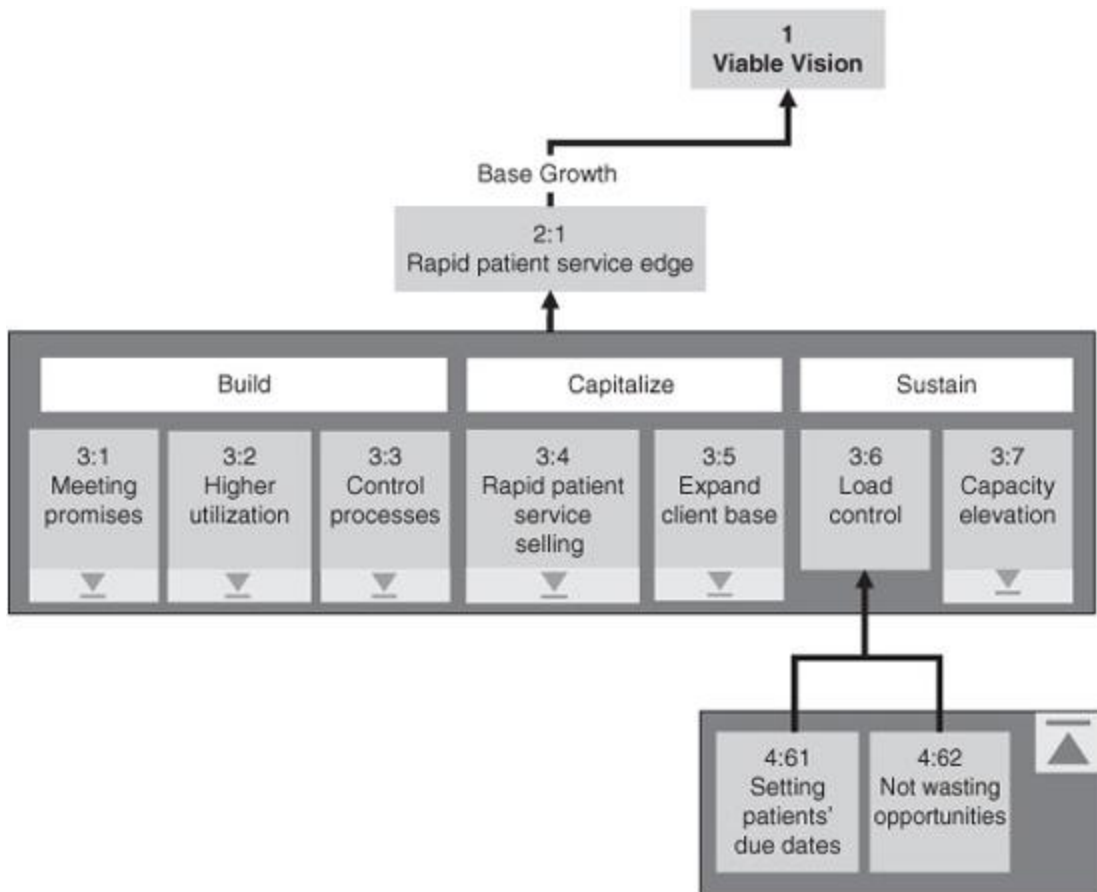


3:5	Expand Client Base
Assumptions Behind Strategy	The number of referrals needed to fulfill the system's Vision is much greater than from past experience.
Strategy	The system is capable of bringing in a growing number of new patients and referral sources.
Assumptions Behind Tactics	There are two excellent sources of referrals—patients and people who know patients (noncompeting doctors, hospitals, specialists, etc.).
Tactic	A staff member is given sufficient dedicated time to follow up with patients, existing referring professionals, and new potential referring professionals to generate a targeted number of new patient leads weekly. This effort is temporarily suspended whenever the lead time to bring a new patient in for his or her first appointment is longer than most new patients would consider acceptable.
Take Note!	When new patient leads grow quickly, it is easy to run out of capacity to service them and result in patient and referring source hostility. New processes of support, control, and measurement are usually needed.

4:51	Lead Generation
Necessary Assumption	When the people in a practice are transaction driven, lead generation is based mainly on opportunism. After a short while, the leads that the practice has are not sufficient to sustain growth for all practitioners.
Strategy	There is a sufficient, constant flow of new patients waiting to enter the practice.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • Many patients who enjoy the benefits of rapid service are willing to provide referrals. • Some noncompeting professionals are willing to refer their clients to a practitioner who can provide a unique level of service. • The characteristics of a person who can extract leads from referrals and identify leads by calls are not the same as the characteristics of a person who can present a treatment plan and gain commitment.
Tactic	Develop and apply a mechanism that requires less and less of the treatment plan staff presenter's time to generate and maintain a constant number of qualified leads.

4:52	Pipeline Management
Necessary Assumption	A practice that is used to dealing with only a few new patients at a time is not set to deal with a quantum leap in numbers of opportunities. Wasting, due to lack of proper attention, a patient or source of referral that had already expressed a genuine interest, is a crime.
Strategy	The best opportunities are not lost due to improper attention.
Assumptions Behind Tactics	When a resource handles too many opportunities, "bad multitasking" is unavoidable.

Tactic	Develop and apply a mechanism to: <ul style="list-style-type: none"> • Define the number of opportunities that the practice staff can handle at a time. • Monitor and prioritize opportunities according to the duration of the opportunities in the sales pipeline (duration in each step and overall duration). • Identify major causes for delays/drop-outs and take corrective actions (many times waiting for first appointment is <i>the</i> major cause of delay). • Monitor the effectiveness of the offer in the various market segments/ product categories to redirect marketing/sales.
--------	--

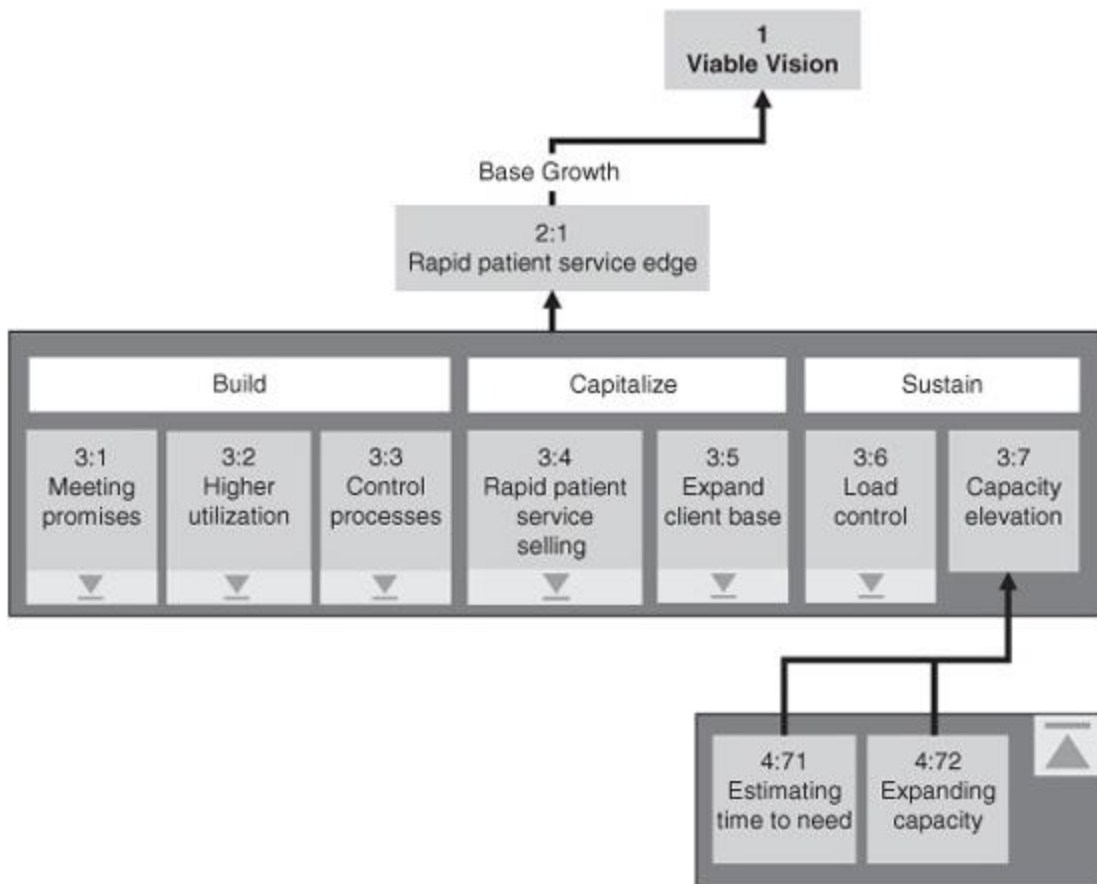


3:6	Workload Control
Assumptions Behind Strategy	When more people are seeking services, it puts additional load on the resources. The patients may have to wait a long time to get into the practice.
Strategy	Predictable lead times, wait times, and quality of service are sustained, <i>irrespective of the growth in practice.</i>
Assumptions Behind Tactics	<ul style="list-style-type: none"> • It is relatively easy to meet all patient time and quality of service needs when the commitments are given based on the workload on existing critical resources and new patients are accepted and staggered according to doctor capacity. • Staff training, cross training, and buffer staffing is in place. • Given enough warning, it is feasible to train or add suitable resources.

Tactic	The mechanism is in place to enable bringing new patients into the practice based upon load on the doctors. The new patient selection and acceptance process is strictly obeyed, even if it means losing some new patients.
Take Note!	When answering a new challenge, it is best to do it with minimum change to the already established practice.

4:61	Setting Patient's Due Dates
Assumptions Behind Strategy	<ul style="list-style-type: none"> As sales grow, the changes in the type of work being done and the blocks of time needed could endanger completion date promises. When sales grow substantially, permanent CCRs appear. If the practice staff continue to commit to promise dates according to fixed lead time, the chances of meeting completion promise dates diminishes.
Strategy	Promise completion dates given by customer service are always met.
Assumptions Behind Tactics	Most of the time required to perform a process is with the patient waiting to be processed by the doctor. Therefore, to ensure continued rapid completion of work, schedules must set aside a percentage of doctor's time for existing patient work and quick initial appointments for referrals.
Tactic	<ul style="list-style-type: none"> The scheduler sets aside a percentage of doctor billable time weekly for treatment plan and referral work based on the past 6 months of actual breakdowns of percentages of treatment plan and referral work. If this time is not booked by treatment plan or referral work 2 weeks before scheduled, it is opened for other work. Due date commitments are given as a week ending date, according to doctor hours per week capacity. Customer service is trained to call the scheduler who gives the shortest promise date according to the available capacity.

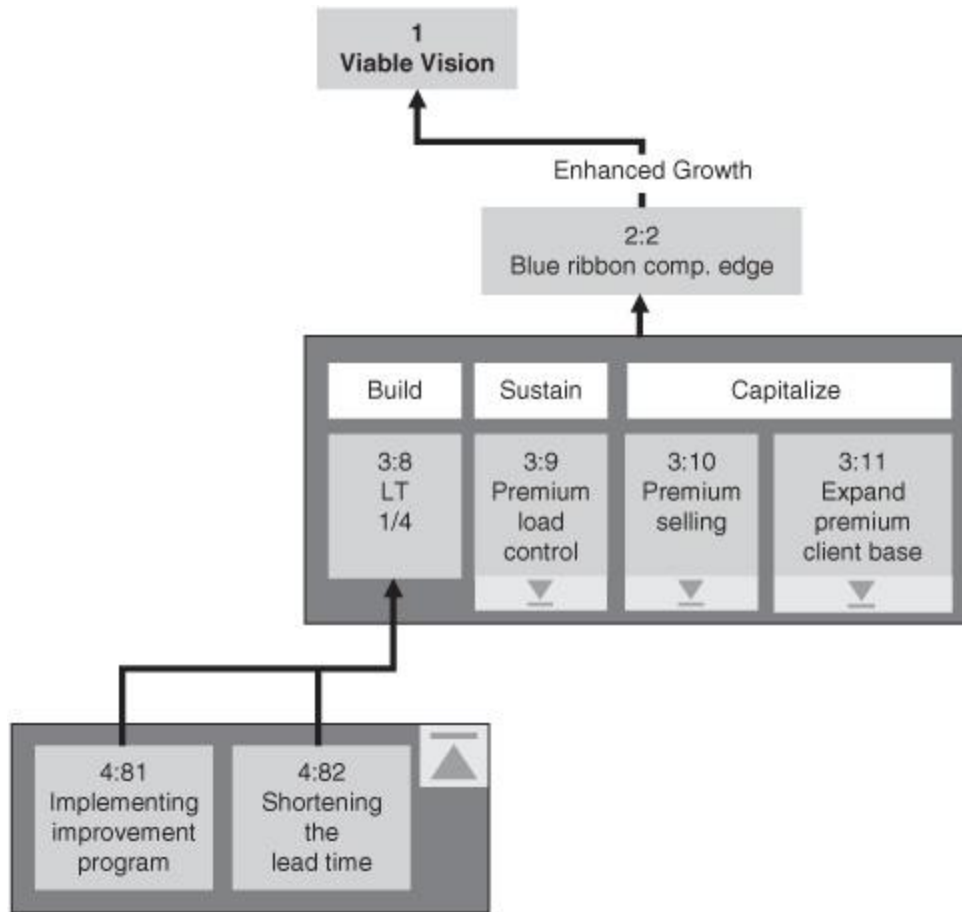
4:62	Not Wasting Opportunities
Assumptions Behind Strategy	Giving due dates based on doctor load may result in very short lead times when doctors are underutilized. Giving something for free jeopardizes the ability to charge for it.
Strategy	The practice does not waste the opportunity to command high premiums for shorter lead times.
Assumptions Behind Tactics	<p>The way to achieve all the following requirements:</p> <ul style="list-style-type: none"> Synchronize due-date commitments with available doctor capacity. Not give (for free) commitments that have shorter lead time than what patients would typically be quoted from other doctors. One mechanism for scheduling and controlling, . . . is to quote no less than a minimum duration, even when the doctor capacity exists to complete the work more quickly.
Tactic	<p>The due date is committed to be the longer of:</p> <ol style="list-style-type: none"> Typical doctor quoted duration (standard duration), or Duration according to the doctor's capacity. <p>Patients continue to be scheduled using standard duration.</p>



3:7	Capacity Elevation
Assumptions Behind Strategy	When total lead times are too long, some of the patients and referring doctors may be lost. The practice's growth and lead times may be limited by staff available in local markets.
Strategy	Desired patients are not lost due to service durations that are too long or due to the inability to expand capacity due to staff availability.
Assumptions Behind Tactics	Profits increase when additional sales are gained for just an increase in staff. After some time, the first actions toward the Vision bring the practice to be cash rich. At that stage, the load of added investment in additional location, space, and hiring another doctor is not a barrier.
Tactic	A mechanism is in place to open rapidly the capacity (staff and space) to prevent significant revenue loss caused by long wait time in the pipeline (patients seeking other doctors).
Take Note!	Too often, capacity expansions resemble playing Russian roulette (making large long-term commitments too late or too soon based on vague knowledge of probability, amount, and timing of need).

4:71	Estimating Time To Need
Assumptions Behind Strategy	Not knowing when additional capacity will be needed leads to increasing expenses/investments too early or (even worse) too late.
Strategy	The practice schedule has a good enough evaluation of the time left until the committed lead-times will start to be too long.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • The practice starts to run the risk of jeopardizing sales (entering into the "danger zone") when the promised completion date starts to be longer than typical quoted durations. • The time until the practice enters the "danger zone" depends on the pace at which the front of the load on the doctors is advancing (and expected to continue advancing).
Tactic	The practice implements a mechanism that constantly analyzes the pace at which the front of the load on the doctors advances (increase in billable utilization per week, starting 4 weeks out). From this mechanism, the practice derives a reliable prediction of the time until they will reach the "danger zone."

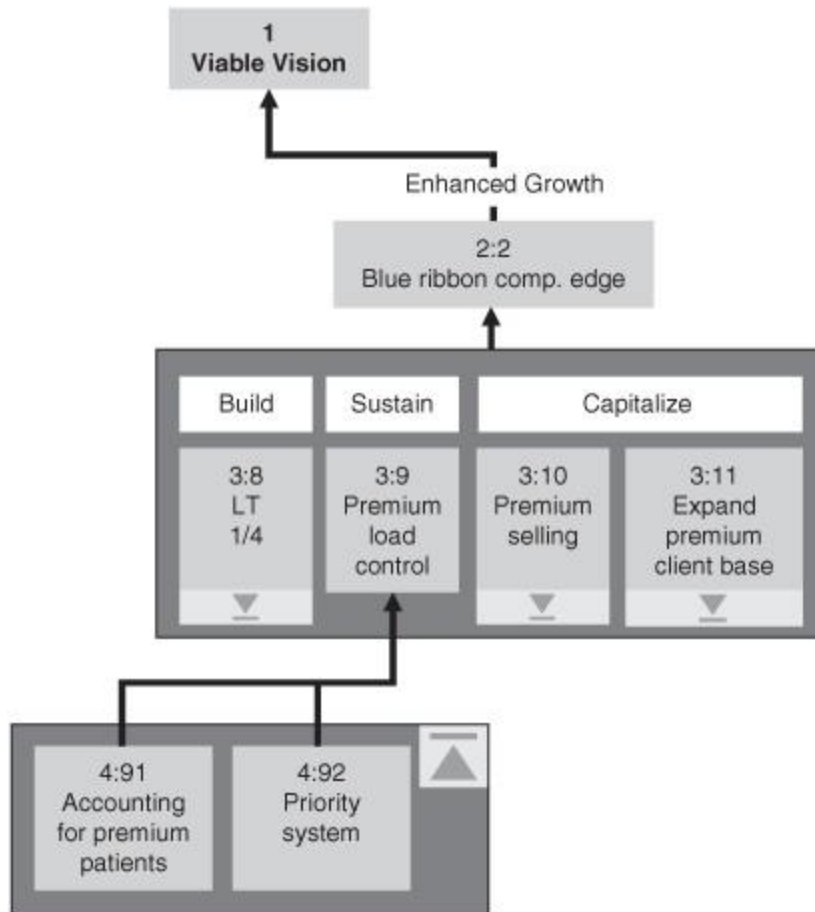
4:72	Expanding Capacity
Assumptions Behind Strategy	<ul style="list-style-type: none"> • Not knowing how much time it will take to have additional capacity leads to increasing expenses/investments too early or (even worse) too late. • The time from making the decision to open capacity until the additional capacity is available is heavily dependent on the level of preparations (actions that can be taken without any final commitment).
Strategy	Capacity expansions are done in time to prevent damage to patients/referring doctors.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • The knowledge of what type and amount of capacity is needed for the next expansion step is available when operations are run by Constraint Management combined with Buffer Management. • The time and needed preparations to add capacity depend on the type of resources needed. • When proper preparations are done, the time from decision to having the additional capacity available is well known.
Tactic	<ul style="list-style-type: none"> • The system builds a team in charge of capacity elevation for equipment, space, and people in all functions. • The capacity elevation team has a monthly capacity elevation plan ready for execution and approved by the practice owners. • The stakeholders agree to hiring and training people and buying equipment in time to meet market demand advance indicators.



3:8	Premium Service Turnaround Time
Assumptions Behind Strategy	If a practice is constantly expediting treatment plans, the scheduling and the practice are often in chaos.
Strategy	Dentist's treatment lead times are reduced significantly.
Assumptions Behind Tactics	When improvements in each area are guided by a combination of TOC, Lean, and Six Sigma (TLS), the treatment lead time can be cut to less than half without compromising on quality of patient care.
Tactic	TOC, Lean, and Six Sigma skills are brought to the employee level on a broad scale. Broad-scale improvement programs are set and constantly guided by the one or two factors that have the biggest impact on doctor time and patient satisfaction.
Take Note!	To shrink the lead time, not only should the reasons for delays be removed, but also the mechanism that releases the patients into the system should be adjusted accordingly.

4:81	Implementing Improvement Program
Assumptions Behind Strategy	Most local improvement initiatives that use good tools (TOC cause-and-effect analysis, Lean, and Six Sigma techniques) do improve the local performance but often those local improvements do not translate into global improvements.
Strategy	All local improvement initiatives do contribute meaningfully to reducing lead time and increasing capacity.
Assumptions Behind Tactics	Recording the reason a patient is in the yellow or red zone of their treatment plan buffer (recording what the patient is waiting for), and analyzing the frequency of patients "waiting" for the same reason (BM Pareto analysis), is a prudent way to identify where an improvement initiative will contribute meaningfully to overall performance (especially to shortening lead time). Knowing the approximate touch times also provides good intuition on abnormal treatment plan lead times.
Tactic	Implements local improvement programs, which are guided by Buffer Management and touch time analysis. Events putting patients into yellow or red zones (equipment maintenance, poor skills, rework, poor supervision) are quickly addressed.

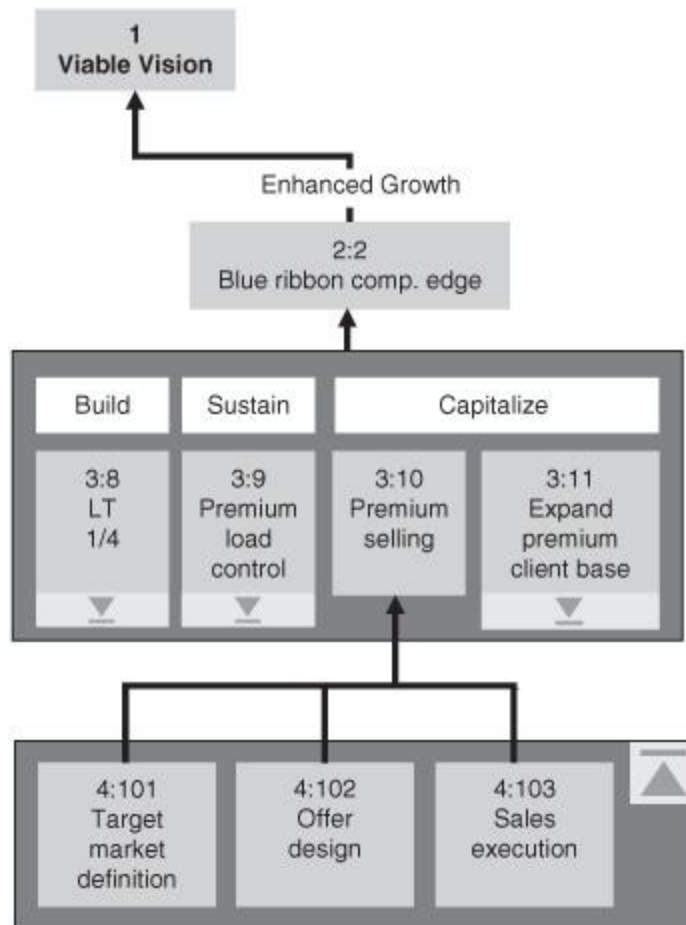
4:82	Shortening the Lead Time
Assumptions Behind Strategy	The lead time of a single treatment plan is partly determined by how many patients are also in process at the same time.
Strategy	The duration of all patient appointments is shortened.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • Shortening the appointment time requires shortening the practitioner's time to do his or her portion of the work. • The buffer needed to ensure rapid patient completion on time is mainly to accommodate variability in touch times. • Some variability can be reduced.
Tactic	When less than 5% of the appointments penetrate the red zone of the buffer, the appointment buffer is reduced.



3:9	Premium Load Control
Assumptions Behind Strategy	Shrinking individual appointment lead times is not sufficient to ensure accelerated service (a major portion of the time is the time until the patient is seen in the office).
Strategy	Doctor has the ability to treat a considerable portion of the practice volume in less than 10 days, including time for emergency patients.
Assumptions Behind Tactics	<ul style="list-style-type: none"> • When a patient jumps the queue, it disrupts the care of other patients unless capacity was allocated in advance for such events. • For those patients who jump the queue of regular patients, total treatment lead time is equal to appointment lead time plus inter-appointment time. • For patients who get top priority in the practice, total treatment lead time is much less than normal treatment plan time.
Tactic	Reserve enough capacity for premium patients (when giving a due date for a regular patient, the fact that capacity is reserved for premium patients is taken into account).
Take Note!	Dealing simultaneously with two vastly different types of patients (regular and premium) can severely complicate the practice.

4:91	Accounting for Premium Patients
Assumptions Behind Strategy	Premium sales require short service times and therefore cannot wait for the next available capacity slot.
Strategy	The system has the ability to capture enough premium sales opportunities.
Assumptions Behind Tactics	When the capacity allocated to premium is not used for premium patients, it still can be used for regular patients. Therefore, the system can be liberal in its estimation of the amount that is allocated to premium.
Tactic	System allocates some of the capacity for premium patient appointments and gradually increases the amount allocated. (The balance of the capacity is used to accommodate the regular patients.)

4:92	Priority System
Assumptions Behind Strategy	The need to deliver some patient treatments in a short range of lead time may create a complicated priority system in the office.
Strategy	The office has one simple and robust priority system.
Assumptions Behind Tactics	At the time of their release, patients using the same priority system automatically will be given high priority when they need to be finished in a shorter time.
Tactic	The regular priority system is the <i>only</i> priority system in the practice.

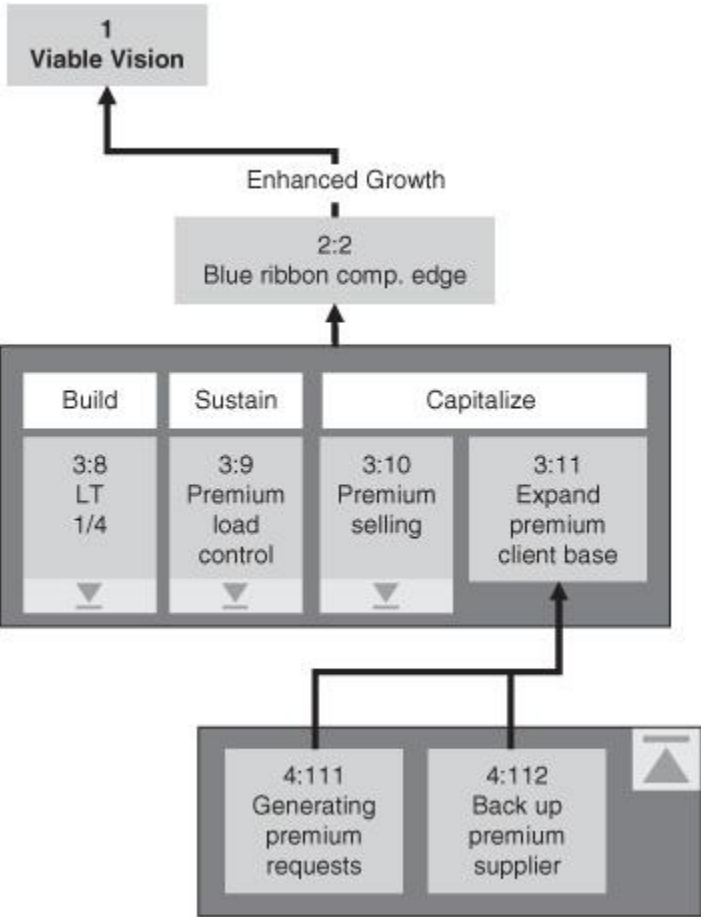


22 4:103 is not included.

3:10	Premium Selling
Assumptions Behind Strategy	Health care practices do not know how to sell services for substantially higher prices than "regular pricing."
Strategy	Practice front line and patient care coordinators are proficient at selling the premium offer.
Assumptions Behind Tactics	<p>The constant pressure of the market/HMOs to reduce prices causes health care providers and their staff to be very skeptical about the feasibility of getting premiums.</p> <p>When a front line staff or doctor is uncomfortable with the deal, or does not fully understand the benefits to the customer, he or she might jeopardize a sale of premium service to patients.</p> <p>When staff have a successful experience with an "unrealistic" offer, their attitude changes to "Of course patients will pay more for such a high service level" offer.</p>
Tactic	The practice staff is trained in how and when to present the premium offer and are "hand-held" in their first attempts. Success stories are shared weekly with all staff.
Take Note!	In many aspects, preparations and training done for one offer (Rapid Patient Service) are not adequate for a different offer (Premium).

4:101	Premium Target Market Definition
Assumptions Behind Strategy	Pursuing wrong prospects is not just a waste of valuable resources (money, sales capacity, time, etc.), but also it can lead to the "conclusion" that the direction is invalid.
Strategy	Front staff know which types of customers are best suited for the premium offer.
Assumptions Behind Tactics	<p>Not all prospects for the rapid patient service offer have a significant need for premium service.</p> <p>There are patients to whom premium service is a significant need; however, they are too risky or require excessive efforts to work with.</p>
Tactic	<p>Target markets are defined according to conditions that are:</p> <ul style="list-style-type: none"> • Easily checked. • Relate to a non-negligible number of prospective patients. <p>The conditions help front staff to prioritize patients according to:</p> <ul style="list-style-type: none"> • The degree to which the patients are willing to pay a premium for rapid service. • The estimate of the ratio efforts/returns. • The degree of business risks.

4:102	Premium Offer Design
Assumptions Behind Strategy	When the details of an offer are not clear, front staff will not sell it because they think the risk is too high or the benefits are trivial. When the details of the offer are not constructed to mitigate risks and ensure benefits (to health systems and patients), the outcome may be losing many good sales opportunities or profit margins.
Strategy	System has a detailed premium offer that guarantees exceptional benefits to patients while ensuring that it is not taking any real risk.
Assumptions Behind Tactics	To construct a good offer, four elements must be thoroughly understood: <ul style="list-style-type: none"> • The net benefit for the system and end customers (target audience) relative to a standard offer. • The benefits to providers of care. • The risk for the target audience (relative to risk they take in a standard offer). • The risk to doctors (relative to the existing risk doctors experience in a standard offer). Ensuring the benefits provides the detailed backbone of the offer. Mitigating the above risks provides important details of the offer.
Tactic	All of the health care staff build the details of the premium offer (penalties, pricing, lead times, and terms and conditions), maximizing the benefits to doctors and patients while minimizing the risks (to both the patient and the doctors).



3:11	Expand Premium Client Base
Assumptions Behind Strategy	<p>Servicing all of the patients (regular as well as premium) unnecessarily limits the ability to capitalize on premium patients because:</p> <ul style="list-style-type: none"> • The number of emergencies existing in the market dwarf the capacity of the system to fulfill them. However, the fact that there is a need in the market and the fact that there is someone who can fulfill the need does not yet guarantee sales.
Strategy	Over 20% of the system's volume is sold at premium prices.
Assumptions Behind Tactics	It is possible to obtain premium sales exclusively from some referral sources when no other health care provider can do the same.
Tactic	Health system launches a wide-based (well-manned and managed) program to ensure that enough of its potential market is aware of system's premium service.
Take Note!	Marketing is essential, but there are environments where even the best marketing is not sufficient.

4:111	Generating Premium Requests
Assumptions Behind Strategy	Many patients do not think of a specific provider when there is a customer with urgent needs.
Strategy	Practice X's premium offer is the first thought that comes to enough referral sources', patients', and friends' minds when they need urgent care.
Assumptions Behind Tactics	The general need for fast response service is big and continuous. Prudent marketing is bound to yield fruit, where a push for an immediate sale will usually fail.
Tactic	Create a team to identify appropriate marketing and sales channels and launch a geographically expanded marketing campaign to brand Practice X as <i>the</i> rapid care provider.

4:112	Back-Up Premium Supplier
Assumptions Behind Strategy	<ul style="list-style-type: none"> • Servicing all of the patients (with regular as well as premium service) limits the ability to capitalize on premium patients. • There are environments where the best marketing campaign will not be sufficient because the time to get approval of the whole value chain is very long. An example is multidisciplinary care patients.
Strategy	The share of premium patients is constantly growing.
Assumptions Behind Tactics	In many cases, the high premiums justify the investment needed.
Tactic	Practice X invests in becoming a backup premium service care provider to a growing community of doctors.

Addendum: Excerpt from the Book *Vision for Successful Dental Practice* by Gerry Kendall and Gary Wadhwa

Steps to success for a private, academic, or government-run dental practice

1. Set a clear goal for the practice. It could be 100 percent increase in the value of the practice or profits in 4 years. Academic, government-run non-profit organizations can have a goal of 100 percent increase in patients served in 4 years while maintaining high quality and low cost.
2. Use a performance measurement system that captures the system performance rather than the individual performance of a department or particular doctor. TA and Finance focus on overall system performance.
 - a. Net Profits (NP) = Throughput (T) - Operating Expenses (OE); Throughput is the payment in the bank after completing the expected treatment on a patient.
 - b. Investment (I) decisions must filter through this formula. Investments are done in order to provide services to the patients or, in other words, to improve Throughput (T). If $T > OE$, it is a good investment because it will result in higher profits. Investments require capital and interest payments over a specific time period like 10 to 15 years. Some investments depreciate faster than others do. All investments result in increases in OE over time. The intended purpose of investment is to increase T and this increase must be greater than the OE due to this investment.
 - c. Return on Investment (ROI) = NP/I (Investment). Investment must be considered over the time period.
 - d. All marketing and advertising decisions must increase $T >$ cost of marketing and advertisements.
 - e. All expansion of physical location, addition of operatories, purchase of equipment, and offering of

specialized services must go through the tests of TA.

3. TOC's basic premise is that every complex system is easy to manage (inherent simplicity) and it usually has one constraint or weakest link. This constraint determines the productivity or Throughput of the practice. Dental practices ideally should have the doctor as the key constraint, but sometimes the constraint could be an x-ray or CT scan machine or microscope in an endodontist's office or limited physical space like in metropolitan cities where the available space is limited and extremely expensive. If the doctor is sitting idle, the constraint is assumed to be in the marketplace, which means that the practitioner might not be attracting patients to the office, or the constraint is internal and is obstructing the flow of the patients through the key constraining resource, the doctor.

It is usually easy to map out the different steps that the patient has to go through in our system in order to get dental care. We can then approximate the time it takes at each step and the usual delays in the flow of patients through these steps. This can give us an overview of where the constraint is located. If the patient has to go to an orthodontist, periodontist, and endodontist prior to completing the treatment, we could assume that the orthodontist's office will be the key constraint because it takes the longest to complete the orthodontic treatment. We might be surprised that sometimes the wait time to see an endodontist might be 3 months. The whole treatment might take 1 week but the total time to process the patient through the endodontist's office is 3 months and 1 week. This might be the key constraint to completing the patient treatment. The unfinished

treatment is not Throughput until the whole service is completed to the patient's satisfaction.

4. Determine how to exploit this constraint. We focus on the means to make our constraint both effective and efficient. Let us assume the constraint in our practice is the doctor time. We have to make effective and efficient use of the doctor time. Effectiveness means a deliberate action of focusing on the correct product mix segments by servicing only the select group of patients. Efficient means that number of patients seen in a given time increases without affecting the quality of service. Frequently decisions about effectiveness are made based upon fees for a particular procedure like a dental implant. Most of the time various other factors, like precious doctor time, investment costs, truly variable costs, and opportunity cost where we must forgo doing restorative work in order to do this implant procedure, are not considered in decision making. The formula to make a decision of product mix (which procedures to focus on) is usually simple: $\text{Throughput (T)}/\text{Constraint Unit or Doctor Time}$. When comparing one group of procedures versus others or when considering referring patients to specialist versus learning to do the procedure in house, the above formula will help make the decisions. The efficiency of doctor time increases when we are careful about not wasting the precious doctor time on tasks unrelated to patient care.
 - a. Total kit concept where everything is ready for the doctor to treat the patients. This includes equipment, tools, laboratory work, radiographs, and information about the patient from other specialists or practitioners including the physician's medical clearance, if necessary.

- b. Equipment and tools have preventive maintenance programs so that there is no surprise breakdown of equipment.
- c. A standardized flow sheet similar to Basic Life Support or Advance Cardiac Life Support that outlines all the treatment steps is used. This helps the whole team know exactly what is expected in the next step.
- d. Workplace organization ensures that everything has a place and everything is in its place.
- e. Supplies are always available when the doctor is working; they never expire or run out. On the other hand, the supplies are not ordered excessively because this will increase OE.
- f. Emergency equipment and supplies are always updated and checked on a periodic basis.
- g. Workplace is meticulously clean and welcomes everyone to come to work.
- h. Quality of work is important because time is wasted in redoing the procedures instead of doing a new procedure that we could have done.
- i. Health care has a lot of surprises like patients coming late or arriving early, patient's expectations change, patient and staff personality and communication styles, and the procedure can have some unexpected delays or complications. The staff, who helps the doctors offload their work, might be absent or unreliable. If the staff changes the new staff might not have the requisite skills. It is important to have protective capacity (capacity to accommodate Murphy and maintain patient flow) of the staff to ensure that the doctor time is never wasted. The protective capacity is the extra skill sets or extra staff, who might at times appear to be standing around, but they actually help to protect the precious doctor time.

5. Subordinate everything to the above decision: The challenge is to control this environment where every patient is unique and there is never a predictable time to get things done.
 - a. Understanding that there are two goals, one to protect the doctor time and second to ensure that the patient is not unnecessarily waiting too long, which can result in the patient's dissatisfaction and through blockage can cause loss of throughput.
 - b. The time prior to the patient seeing the doctor is considered the doctor buffer. The patient might be present prior to the doctor finishing the proceeding patient. Most dentists work at least two or three chairs. This means that the time it takes to register the patient, take an x-ray, or have laboratory work ready for the patient must be done and the patient is ready in the second chair prior to the dentist finishing the first patient. If the dentist takes a long time to complete the first patient, there must be a signaling system to inform the check-in person so that the staff does not make the patient wait in the operatory unnecessarily. If more than two procedures took longer than planned and a long wait will result, then the staff must have a system of informing the patients regarding this delay. The flow manager keeps the patients, who have arrived at the practice, occupied with coffee, tea, magazines, TV or Internet in the waiting room, The flow manager admits the patient into the system only when the doctor catches up. This prevents the staff from multitasking and prevents the staff from being tied to a patient when no work is being done.
 - c. Since every patient is different, he or she could take different amounts of time to complete treatment. This environment is similar to the multiple projects environment. We must prioritize and have some

computer calculate the staff utilization. The flow manager directs staff to different workstations as the need changes. Such software is not available at this time for health care applications; however, it is being developed.

- d. Buffer Management helps us identify where and why the delays occur. If most of the delays are due to doctors not starting on time, we can figure out how to influence the behavior of doctors. If the patients always arrive late, we can start reminding them to come 15 minutes earlier.
6. Elevate the constraint: Once the company has fully exploited the doctor time and has subordinated everything else to doctor time, it is time to elevate the constraint to the market by hiring another doctor. The market becomes the key constraint. We can start the same focusing principles to determine how to exploit the market.
7. The last step is to ensure that inertia does not set in that could otherwise become a constraint. Once the practice is doing well, everyone becomes relaxed and happy with his or her achievements. The processes and systems start to slip and they start to go out of control, which results in a downward spiral. Be aware of this tendency.

CHAPTER 32

TOC for Large-Scale Healthcare Systems

Julie Wright

A patient opens a consultation with a doctor by saying “Doctor, it hurts when I do this.”

The doctor asks “Why do you have to do ‘this?’” mimicking the patient.

“Because I have to achieve that,” replies the patient moving around the room.

“OK, but what if you could achieve ‘that’ by doing ‘this’ differently?”

“That would work,” said the patient excitedly.

“So you agree.... If ‘this’ hurts it’s best to stop doing it? You’ll have time to heal and you will still get the results you need?”

“Of course.... Thanks doc!”

Introduction

Unlike the practice of medicine by individual physicians, the practice of medicine within large-scale healthcare systems is a relatively new phenomenon. As industrialization concentrated populations in urban areas, medicine followed

suit and began to be practiced by groups of physicians. As the collective provision of medical services flourished, exponential advances in the diagnosis and treatment of patients caused the medical profession to divide into specialties. Now it is often the case that a patient's care episode is dependent on the services of more than one specialist to reach a successful conclusion. Because of the division of specialties, patients are often forced to interact with many different people and services to secure the holistic treatment their conditions dictate.

While the delivery of healthcare is moving toward a more holistic model, the infrastructures within which it is delivered is still, for the most part, segmented with patients often being required to tolerate unnecessary waiting times between the receipt of separate services.

This chapter aims to show what needs to be changed, what the systems need to strive to achieve, and how to begin to identify the causes of these delays and to eradicate them, eventually redesigning the delivery of services to fit the ability of patients to absorb treatments, not their tolerance to waiting.

By successfully melding the many diverse services with a systemic approach, it is possible to increase the capacity of existing systems, reduce overall costs, and improve the quality of patient care along with the working environment of the people dedicated to the profession. How to stop doing what hurts and replace the current actions and behaviors with better and more effective practices that will benefit all is the goal.

Copyright © 2010 by Julie Wright.

Why Change

Why Healthcare Systems Need to Improve

If you look hard enough, it is possible to find large-scale healthcare organizations in almost every shape, size, and form imaginable. To be considered a large-scale healthcare provider, an organization should be able to treat a wide spectrum of human conditions ranging between prevention, disease, and accidental trauma and almost all conditions in-between. Some patients' conditions need immediate urgent care, others longer term treatments. Some systems may also provide additional specialist services from assisted conception to palliative care and from rectifying congenital defects to, perhaps, the genetic engineering of zygotes.

The locations for the delivery of care can vary widely as well; from urban based, high-tech, tertiary, multispecialist facilities to rural lone practitioner family physicians, they all contribute to the larger scale healthcare systems that we will all find ourselves entering and using during our lives.

Some hospitals have Emergency Rooms (ERs) and some do not. They come in different sizes to suit the needs of the population being served; and some are even licensed to perform different functions for different levels of trauma. Some Emergency Rooms are the front door to the only healthcare facility for miles around and in others the local physician is the Emergency Room surgeon and general practitioner. In some, the general practitioner is the gate-keeper to the hospital's services even though they do not practice there.

There are facilities that focus on providing care for a very narrow range of conditions, such as standalone diagnostic centers and clinics that provide services for non-life threatening, elective surgery.

Around the perimeters, the large-scale healthcare organizations are providers of alternative therapies, some of which are gaining credibility and are being absorbed into the practice of scientific based Western medicine.

A large-scale healthcare organization can be a huge group comprised of hospitals, clinics, pharmacies, transport services, nursing homes, rehab facilities, and diagnostic centers, with large administrative offices that are remote from clinical areas, and some have international operations that are unbounded by international borders or politics. Alternatively, they can be a loose association of any type of clinic or surgery bound by a cooperative willing to support each other and their patients.

They can be not-for-profits organizations run by governments such as military services, charitable bodies, or socialized medicine, or as profit-oriented businesses, or as a variation of every shade of business model between the two ends of the spectrum. Blends of for-profit and not-for-profit can exist and function side by side within a large-scale healthcare organization.

These organizations can have another dimension—some are religious and others are secular; many have educational affiliations, as in teaching hospitals attached to medical schools, or are entirely privately owned and operated.

Whatever mix of services and provisions an individual healthcare service system possesses they all have a common need, that of being able to generate or acquire sufficient cash to operate. Even highly motivated not-for-profit providers have no mission if they do not have an operating margin. For-profit and not-for-profit organizations cannot operate at a financial loss no matter what their source of income: fees, donations, endowments, etc. In many countries, not-for-profit organizations qualify for favorable tax breaks. The other main difference between for-profit and not-for-profit operations is that the profit generated by not-for-profits is not paid out to shareholders as a dividend, as for-profits are obliged to do, but the profit or margin is used to sustain and grow the organization. Therefore, a not-for-profit organization, or perhaps “not-for-dividend” would be a more accurate description, will be

unable to fulfill its mission if it is unable to generate a profit, or a margin.

Therefore, no matter what the size, shape, location, mission, or orientation of large-scale healthcare organization, they are all facing all or a mix of the same problems:

- Growing and aging populations—this translates to more patients and a growing demand from the same population.
- Less money—as demands for better value, quality, and quantity increase for the same or lesser amount of healthcare spending.
- More technology—to keep pace with advances in the field of medicine and its administration.
- Higher expectations—of a continually better-educated consumer in most part due to access to medical information via the Internet.
- Increasing competition—especially in more developed societies.
- A need to provide new medical services to currently underserved populations.
- An insufficient supply of clinicians—both physicians and nurses are in short supply globally.
- A more mobile population capable of spreading disease faster than ever before.

Healthcare is an industry that will never lose its client base as long as our race survives and it is one of the most regulated, if not the most regulated, industry in the world. It employs the best-educated workforce in the world and in some cases offers the highest and lowest salaries of any profession.

In short, large-scale healthcare organizations can be as difficult to categorize as we are ourselves and their

problems can be as diverse as diseases we can suffer from.

Given the huge range of diversity present in healthcare systems, the only accurate model that can be drawn of large-scale healthcare systems is that of a black box into which people enter as patients and from which they leave, with a wide degree of altered states of being from a clean bill of health to dead.

The Goal of Healthcare

The human race has an insatiable appetite for healthcare. This appetite reaches beyond treatment well into the realms of prevention. It is commonly acknowledged that “prevention is better than cure”—when it can be achieved. Inoculation and healthy living practices improve life expectancy, but thus far rarely in enough numbers to release the clinical capacity to treat all of those who need care. No matter what the mode of delivery, socialized or private, there are still sectors of every population that can benefit from additional professional healthcare globally. Therefore, every large-scale healthcare system needs additional capacity to treat more patients.

Medical technology continues to advance, in many cases, quicker than the delivery system can bring the advances to the patients. The advent of the Internet has given the public unprecedented access to news about new treatments and online diagnostic tools and medical Websites are educating patients far more than ever before. The expectations being placed on the medical profession are the highest they have ever experienced and will be unlikely to decelerate in the near future. The healthcare industry is under tremendous pressure to treat patients better to achieve more effective results than in the past.

In the practice of medicine, time is often of the essence. The need for the immediate treatment of trauma is often

well provisioned, but even in the most developed of societies ERs get backed up as the not always predictable ebb and flow of patients present themselves for treatment. In contrast to this, the advances made in early detection, more accurate diagnoses, and more effective treatments of less acute but more long-term, chronic conditions has, in some cases, exponentially increased the numbers of people needing lifelong treatment, support, and medication. With the rise in expectations, there is a reduction in the tolerance of the time people are prepared to wait for a medical consultation. There is a pressing need to treat patients sooner than in the past.

The Internet has also given the public access to healthcare performance data. Through measurement and benchmarking, many areas of medicine today are open to scrutiny by their consumers. Choice, even in some socialized healthcare systems, is fast becoming a perceived right of the healthcare consumer around the world. In some countries, health-care is considered to be a basic human right that carries with it statutory legal rights for the individual.

With good physicians and a high proportion of facilities experiencing increases in demand and with some poorly performing services struggling to attract patients, it is an imperative of healthcare providers to keep improving both now and in the future.

Therefore, the global goal of healthcare is to be able to *treat more patients, better, sooner, now and in the future.*

What to Change

Where to Start: Government or Facility?

There are many opinions on how healthcare should be funded and who should be responsible for its delivery.

Socialized healthcare has much to commend it as well as to condemn it. The exact same thing can be said of privatized medicine. The only consensus that can be reasonably reached about the best mechanism for managing the funding and management of healthcare is that there is currently no one best way and each of the methods used thus far appear to fail any given measure of “value for money.”

Tackling the problems facing healthcare at the governmental level of any country is a long, laborious procedure that all too often results in unsatisfactory compromises.¹ Very few people or organizations work within a sphere of influence large enough to be able to have a meaningful impact on a national or legislative level on the delivery of healthcare. If we, as individuals or organizations, strive to change healthcare from the very top down, through the representatives of our respective governments, we will have a mammoth task on our hands with very little chance of success.

However, those of us working within or consulting with healthcare facilities do have a chance of making a difference. Therefore, we need to recognize the limitations of our sphere of influence and be prepared to work within it.

Unlike industry, healthcare is an industry sector that is, for the most part, prepared to share best practices and ideas and processes for improvement because it recognizes the need, even between competing facilities, to contribute to the common goal of trying to treat more patients, better, sooner, now and in the future. This openness supports the numerous journals and publications covering medical advances and the management of healthcare. Healthcare openly admits its need to improve and it is prepared to consider and share ideas and processes that will help it to

improve. However, there are far too many instances when the “silver bullet” for one system or facility is adopted by another facility without fully understanding why it was able to be so successful in the first place. Because of the driving need of healthcare managers and administrators to improve the performance of their facilities, many of them have fallen prey to consultancies and methodologies that:

- Do not address their core problem and therefore fail to achieve the operational improvements achieved in other facilities.
- Fail to yield an effective return on investment.
- Are strangely familiar to longer serving staff who claim to have “seen it all before.”
- Fail to take into consideration the concerns and reservations of the people who are expected to implement the changes.

However, these experiences have failed to quell the intuition of the industry that there must be a better way to manage these systems and produce better results.

That intuition provides the imperative for facilities and systems to continue to seek out, adapt, and adopt new improvement methodologies.

As with all purchases, the caveat needs to be “buyers beware,” unless the facility or system is able to prove to itself that it knows what its core problem is, the underlying reason that most (around 70 percent) of its symptoms exist, and that the proposed solution will address them, they will be introducing a fix that will only improve a small proportion of the system and quite possibly an area where improvements will generate more problems in other areas. Some typical examples of these behaviors are:

- Deciding to improve the ability of an *operating room* (OR) suite to process more patients in a facility that does not have sufficient ICU staff to take care of the patients postoperatively. This results in unsafe staffing levels and additional staff being called in to work at short notice, at additional expenses to the facility.
- Deciding to improve Throughput in the emergency room while ignoring the needs of the discharge process. This in turn results in extended boarding (patients waiting on gurneys) in the department because there are no vacant beds to move them into.
- The decision by an entire board to adopt a method to manage waiting lists that proved effective in reducing waiting times in a facility that was operating at 65 percent capacity and in a facility that was already working at 95 percent of its capacity. The result was a very costly program that was unable to deliver the improvements needed because of a lack of capacity, and a group of disgruntled staff who had to find alternate work when the unit downsized.

This propensity to adopt improvement programs with little or no understanding of the systemic effects is not uncommon. However, when systemic improvements that are able to incorporate the differentiated needs of individual facilities are adopted they can produce astounding results; improvements such as increased patient Throughput at levels that far exceed expectations, with little or no increase in resources.

By working at the facility level with the TOC suite of tools, it is possible to differentiate between the core problems of each individual facility, align the staff to be ready to take an active role in systemic improvements and take full advantage of the industry's inherent propensity to spread best practices to other facilities.

The Organic Nature of Healthcare Facilities

Healthcare facilities grow, and sometimes contract, over time in response to local needs and the availability of clinicians.² As facilities secure the services of medical specialists, their infrastructures develop to accommodate the specialists' and their patients' needs. These needs can also change over time as treatment regimes develop and morph clinical offerings to patients.

A good way to demonstrate the organic nature of large healthcare systems is to view them from above the facilities. Hospitals, even new ones, undergo ongoing physical development with wings, towers, and additional buildings being added to house evolving services. Unlike production plants, few hospitals can afford the luxury of suspending services while these additions or renovations are constructed because of the need to provide around-the-clock care.

The need to work in a constantly changing environment poses problems for the staff and patients alike. As the physical plant of hospitals change over time, their operational systems also need to adapt and change to support the changing mix of clinical specialists and new treatment regimens. All too often the number of changes taking place in a single facility at any one time are too numerous to track effectively. This is especially true in facilities that possess a strong silo management culture; one where the predominant mode of management is departmental, vertical, and hierarchical. This form of management has evolved in most healthcare facilities and it is widely accepted that these organizations are too big to manage systemically.

The Human “Engine of Healthcare”

Over the years, industry has been able to automate many processes and, in doing so, increase the accuracy and therefore the constancy and quality of the products it produces. We can now enjoy the exact same products on every continent of the world safe in the knowledge that they will not vary in quality.

While many of the support services utilized within healthcare have benefitted from, and will continue to benefit from, advancing technology, the interaction between a patient and a physician is one area of healthcare where it will not be possible to replace with automated services.

Telemedicine and remote consultations have their place, but these are compromise solutions to provide patients and their caregivers access to a wider community of clinicians. These technological advances are not a substitute for face-to-face consultations that can provide the clinician with a much greater depth of understanding of the patient's condition and therefore their subsequent diagnosis and treatment needs.

Many attempts have been made to standardize the processes that form the interaction between a patient and clinician and while the outcomes of these interactions can share a commonality, the route through different levels of understanding and modes of communication are rarely the same from patient to patient.

The need for effective clinician/patient communication is gaining recognition with medical and nursing schools to the extent that many, if not all, now provide training in patient/clinician communication. In some schools, these programs carry a required pass mark for progression to qualification.

These programs offer evidence of healthcare systems' reliance on the people working within it to provide the people working with patients, and each other, effective means of communication and the ability to adapt the

required communication to a form that will contribute to the most effective clinical outcomes.

In short, successful outcomes associated with healthcare improvement initiatives are much more likely to occur when the “process units” (people responsible for delivering the service) are able to recognize, understand, and resolve relatively simple local problems through the use of standardized critical thinking processes, communication skills, and working practices.

By teaching the staff how to resolve problems effectively and by providing them with commonly understood management taxonomy and subsequent language that can be used on larger, systemic problems, it is possible to achieve greater success than the application of systemic solutions to standardize or improve operational process alone.

The Constantly Evolving Workforce

Being so people-dependent, healthcare has the never-ending task of providing service with a constantly evolving and changing workforce. It is a profession with clearly defined career paths and it has a culture of life-long learning. As additional patient needs are recognized, the necessary scope and depth of learning keeps growing. It is the ability to keep learning that underlies the adaptive nature of healthcare givers and provides the service with its greatest strength.

No matter what the configuration of a hospital, the people working within it can very quickly change the services conducted within its physical confines to meet the needs of their patients. Although this happens unnoticed each day in each facility, this adaptive behavior is most evident in times of large-scale disasters; a facility designed to treat patients with long-term chronic conditions can be transformed into a

triage center for victims; a unit designated to treat children can provide care for adults; a dental hospital can house wounded military personnel. The function of a facility is more dependent on the skills of the people working within it than the physical plant in which they work.

In addition, the ability of staff to adapt to meet the challenges of prevailing circumstances provides the biggest managerial challenge in healthcare today.

As large-scale healthcare systems prepare to treat more patients, better, sooner, now and in the future, they face the task of aligning a workforce that possesses a very wide range of evolving clinical and interpersonal skills to move their organizations forward.

The Reality of Healthcare

In management terms, healthcare is a blend of two types of project management; that of individual patients being “processed” through the system—each patient can be classified as a project because rarely are concurrent patients’ treatment needs identical—and that of operational improvement projects, such as the introduction of electronic patient records, reducing patient waiting times, etc.

The effect of any improvement initiatives, on both of the project management based work streams, should be to generate an overall increase in effectiveness to achieve one or ideally more of the following objectives:

Generate more income, which will provide a facility with more resources to enable it to

- Take care of more patients.
- Offer improved services.
- Reduce patients’ waiting times.
- Continually improve.

When people who spend the majority of their time involved with direct patient care are charged with tasks from the operational improvement projects, they are often being asked to perform what can loosely be described as “extracurricular” activities as often they have to carve time out of their patient care commitments to perform them. When individuals are charged with participating in both types of “project,” they are often placed under tremendous pressure. They feel a loyalty to their patients at the same time as recognizing the need to improve the system that will eventually help them to deliver better or quicker services to their future patients.

Unlike the production environment, the service sector is far more dependent on the people to deliver the desired outcomes needed to succeed. In services like banking, insurance, leisure, and the like, it is possible to standardize many of the processes along the lines of a production environment, a trait that many departments in healthcare facilities mimic. However, the resulting finished “product” is, say, an insurance policy or a vacation; there are inherent mechanisms that can be tweaked or adjusted to meet the consumers’ needs. However, in healthcare the finished product is far less predictable and the producers are far less able to gauge the effectiveness of their efforts as they are often as dependent on the emotive and experiential issues of their patients’ care episode as the science and technology used to satisfy the health needs of their “clients.” It is this human element of both the “production units” and the “raw material” that introduces huge amounts of potential variation into the management of healthcare and which generate core problems³ that can be difficult to predict and even more difficult to generalize without the use of effective analytical tools.

In order to understand fully the generic core problem of healthcare as an industry, it is necessary to dig deep

enough to find a common cause that will take into account all of the wide range of variations presented by a service that delivers very personal care to people.

The core problem of healthcare has to encompass fully the problems experienced by the people as well as the operational processes used within the system.

To verify this statement, one only has to look at the improvement programs that are adopted by seemingly homogenous large-scale healthcare systems represented by countrywide socialized systems.

It is common for overarching socialized healthcare management entities to insist upon the adoption of certain management practices that have brought benefits to a few of their facilities, often through pilot projects.

What these projects almost inevitably overlook is the real current core problem of each facility⁴ and in doing so assumes that because project X worked so well on one facility, its repetition will yield the same results in all of their locations.

What is not established before these projects are initiated is the core problem of each facility and whether the proposed project will eliminate the core problem by rectifying the underlying conflict, or if it will only address lesser problems and symptoms (undesirable effects) being generated by the conflict.

Current Problem Solving Techniques

“Not everything that can be counted counts and not everything that counts can be counted.”

—Albert Einstein

Many healthcare facilities operate a variety of programs to try to gather and address the problems or *undesirable effects*⁵ (UDEs) the staff and patients experience, but few, if

any, are able to effectively conclude this process to the satisfaction of all involved.

Often raising negatives in the form of problems to line managers can be done, but all too often the response from management takes the form of a survey or numeric analysis of data that attempts to quantify the extent of the raised problem and which often leaves the apparently lower ranking problems untouched in the subsequent improvement attempts.

What is often not recognized during these exercises is the degree of impact that some behavioral problems can have on a system. Even if such problems are raised, they are often accepted as being a “fact of life” that has to be tolerated rather than addressed.

Many facilities investigate “adverse events” and treatment effectiveness through the application of a form of cause-and-effect analysis, from which frequently operational changes are implemented based on the findings. This practice is inherent throughout healthcare, in both the medical and operations management fields. These analyses are often the basis of best practice models that are fast becoming the measurements for clinician performance and payment structures. However, the cause-and-effect analyses are all too often only used to analyze exceptional or isolated events and often fail to dig deeply enough to include otherwise unreported negative effects from such incidents—they fail to unearth the deepest root cause of the problem, as pictured in [Fig. 32-1](#).

In addition to this, some facilities even lack an effective way to raise negatives. Some facilities possess cultures that place expectation on their staff to figure out solutions at a local level. Again, these solutions are far removed from the source of the problem.

These modes of problem solving result in the building of operational barriers between departments that serve to

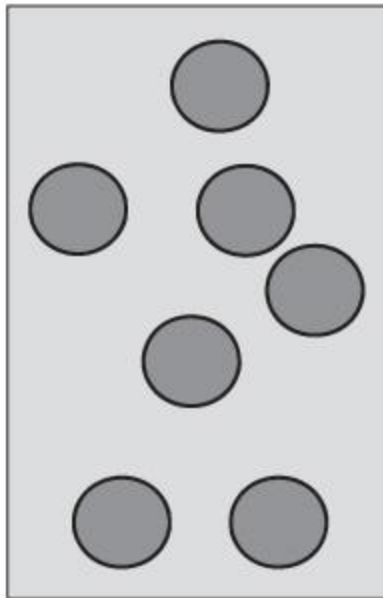
further isolate departments into operational silos and discourage systemic cooperation.

Adapting Industry's Solutions for Healthcare

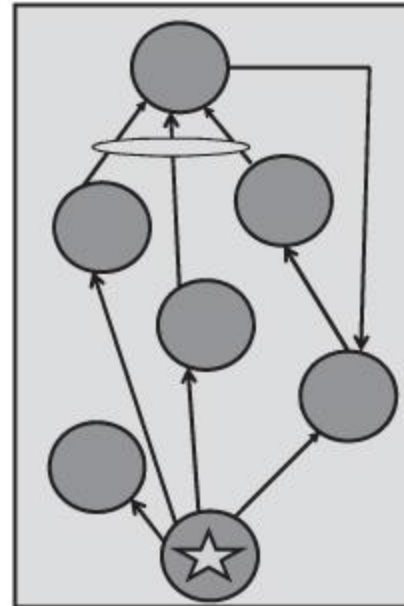
In an effort to try to rationalize the delivery of healthcare, many providers are turning to industry for improvement models. They often view the variation they experience as being a problem that needs to be eradicated, using tools such as Six Sigma and Lean. In some cases, the use of these tools are wholly appropriate *if* the core problem of a facility is a type of constraint that can be effectively addressed using these tools and will contribute to the goal of enabling facilities to treat more patients, better, sooner, both now and in the future. However, if the use of a particular management tool is one that violates *any* of the conditions of the goal or if it demands a compromise solution, then it is not *the* tool the facility needs to use to improve systemically.

Given that healthcare is overwhelmingly people-driven, the majority of the problems that demand the attention of the staff are those generated by the interactions between people. This “noise” has to be greatly reduced before the people working in the system can begin to recognize and be confident enough to address the operational issues that need to be fixed. Until this is achieved, “people will be people” and they will revert to old ways of working of protectionism, watching their backs, and apportioning blame. To reduce the “interpersonal noise” of a system, it is necessary to diagnose why this noise is being generated.

System A: Common practice



System B: TOC



By using TOC cause-and-effect logic tools—as represented by system B—to map the relationships between symptoms, it is possible to focus resources on fixing one, systemic problem at a time, rather than many isolated individual problems.

FIGURE 32-1 Using TOC Logic Tools (© E. M. Goldratt used by permission, all rights reserved. Source: E. M. Goldratt 1999. Viewer Notebook 137.)

Much as a physician uses the presenting symptoms of a patient to make a diagnosis, what is needed to find the core problem of a facility is a rigorous cause-and-effect analysis of the symptoms from which the system is suffering. The symptoms of a system are the UDEs being experienced by the people within the system.

Often the analysis of numerical data will not reveal behavioral symptoms; rather, it will provide a measure of the outcome of the effect of a combination of symptoms, whereas a collection of UDEs in a verbalized form offers insights into behavioral and operational issues that, analyzed with rigorous cause-and-effect methodologies, can be used to expose the core problem that causes them to exist.

To collect sufficient verbalized UDEs to be able to deduce the core problem of a facility, there needs to be a safe environment for the people suffering from the UDEs to voice their concerns. They need guidelines to help them give accurate descriptions of the UDEs, ones that do not place blame onto colleagues, but rather give a clear description of the result of errant actions and processes, and which will not result in future recriminations.

Once a safe platform has been established, it is necessary to make sure that the concerns being raised are addressed in an effective manner.

If both a safe platform and an effective mechanism were in place to understand and address systemic negatives, then far less “interpersonal noise” and fewer operational problems would exist. Therefore, the underlying core problem of healthcare facilities is the *lack of platform and mechanism by which negatives can be effectively raised and addressed* (Wright and King, 2006).

Both a platform and a mechanism are needed because a platform without an effective mechanism to identify and rectify the causes of the UDEs will be ineffective, as is a mechanism that does not address the majority of negatives at a systemic level.

If both an effective platform and mechanism were present in a facility, the UDEs or symptoms being experienced would be of minimal concern and the facility would be able to improve with the following results:

- A minimum amount of disruption to patient care.
- A cooperative workforce.
- The facility working at optimum capacity, generating or securing the maximum possible income.
- Clinical staff would be able to devote almost all of their workday to the treatment of patients.
- Administrative services would be subordinate to clinical services, causing minimal disruption and waiting times

- for patient/clinician interaction.
- A greatly reduced need for clinicians to participate in administrative improvement programs.

What to Change to

Where Should the Constraint Reside in Healthcare?

In an ideal healthcare system, there would be nothing to stop the constrained resource of clinicians from maximizing the time they spend with patients. In fact, the clinicians *need* to be the constraint.

This constraint will never be broken *until* there is enough clinical capacity to treat *all* of the community's patients, with the best available methods, as soon as they need it.

If a facility does have sufficient clinical staff, the constraint needs to be the recovery rate of the patients. Under these circumstances, the only factor that should impede a patient's progress through the caregivers' services should be the patient's ability to heal or recover, with no system or clinician imposed wait times.

These ambitious constraints are far from being onerous; they are the constraints health-care providers and their managers should be striving to establish within their individual facilities.

However, before these ambitious targets can be reached, it is necessary to address the underlying core problem.

Starting an Organization on a Process of Ongoing Improvement

In healthcare, the deepest problem is a lack of platform and mechanism by which negatives can be effectively raised

and addressed is easier to understand in the form of the personal conflict or dilemma being experienced by the people who suffer from it. They are caught in the personal dilemma described in the following Evaporating Cloud⁶ in [Fig. 32-2](#).

This Cloud reads:

In order to [A] treat more patients better, sooner, now and in the future, I need to [B] contribute my expertise to the improvement of our facility; and in order to [B] contribute my expertise to the improvement of our facility, I want [D] to raise reservations about proposed changes.

On the other hand, in order to [A] treat more patients better, sooner, now and in the future, I need [C] not to waste my time (use my time as productively as possible) and in order to [C] use my time as productively as possible, I [D'] don't want to raise reservations about proposed changes.

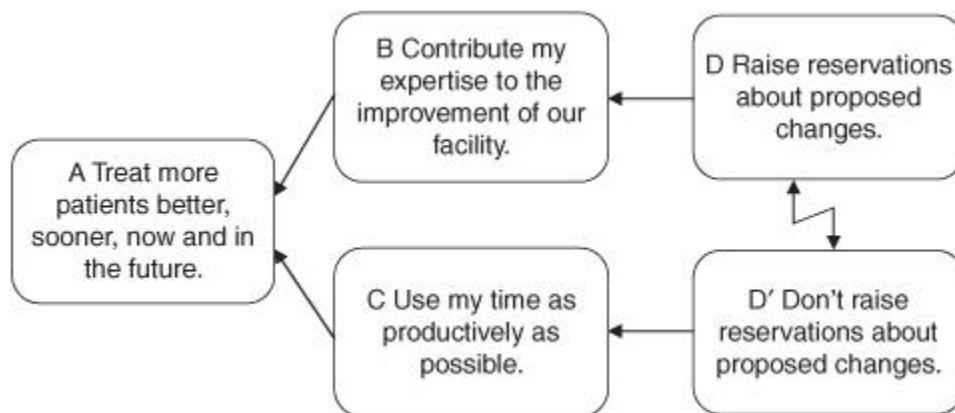


FIGURE 32-2 Contributing expertise.

Obviously, D and D' are in direct conflict, a simple "do or don't do" dilemma.

Some of the assumptions behind some of the arrows of this cloud are as in [Table 32-1](#).

Because this dilemma is so prevalent in healthcare, the people working in it are often unable to prioritize effectively between the demands on their time and they resolve this perpetual dilemma by accepting an ever-increasing administrative workload. This is a practice that, as it becomes more and more common among the workforce, is accepted as being a fact of life in healthcare.

As people feel obligated to accept this fact of life, they try to complete all of the work expected of them and are frequently forced into accepting a compromise solution by working more hours, and taking on more unrecognized and often unpaid tasks and responsibilities than they are contracted for, all too regularly to the detriment of their personal lives.

Conversely, the people who refuse to be put upon are often thought of as being obstructive and uncooperative.

In either case, the inability or unwillingness of these people to raise objections to the proposed change or to suggest alternate solutions is prevalent, and without the ability to resolve this dilemma in a way that does not compromise the essential needs of both B and C, it is not possible to effectively achieve the goal of healthcare to treat more patients, better, sooner, now and in the future.

The top of the Cloud B-D			
In order to [B] Contribute my expertise to the improvement of our facility I want to [D] raise reservations about proposed changes.	Because if I don't raise my reservations the changes will likely have a negative impact on my work and it will be too late to complain.	Because no one else has the knowledge to raise the reservations I have.	Because raising reservations I will be in a position to offer advice as well.
The bottom of the Cloud C-D'			
In order to [C] Use my time as productively as possible I [D'] Don't raise reservations about proposed changes.	Because raising reservations by (email or telephone) will result in an invitation to yet another time-consuming meeting.	Because it will take up too much of my time to attend the meetings where my reservations will need to be voiced to be effective.	Because raising reservations about changes usually ends up with the allocation of more work which I don't have time for.

TABLE 32-1 Assumptions for Raising or Not Raising Reservations about Proposed Change Initiative

This dilemma is generated by the fact that staff is often unable to resolve many of the problems generated by any or a combination of the following:

- Interpersonal conflicts
- Conflicting schedules
- Insufficient resources
- Ineffective operational processes
- Erroneous policies

These problems exist because of the deeper underlying problem of a lack of platform and mechanism by which negatives can be effectively raised and addressed.

Providing a Safe Platform and an Effective Mechanism

TOC offers a number of different ways to identify the core problem of a system, all of which can achieve the same result of arriving at a core problem that, on further analysis, can be expressed as a core conflict.

While it is possible to identify an organizational core conflict through one of the more direct TOC analytical process, such as the Three Cloud⁷ analysis, this process may not provide the breadth of analysis needed to fulfill the need of a system-wide platform to air problems. It is necessary to provide the staff with a broad-brush relational map to show from where their own experiences emanate and how the subsequent behaviors they are forced to exhibit are the result of the way the system is structured and operating.

When the system being analyzed is more heavily dependent on people as opposed to mechanical processes and process units, it is necessary to take a more detailed approach to finding the core problem.

At this stage of the process, the staff has neither the time nor the skills to identify their own facility's core problem; therefore, it is necessary to provide both the platform and mechanism for them.

Each facility's personnel must understand what it needs to change⁸ before plans can be made to facilitate change. Few people, if any, in any given facility have a clear picture of how their current activities generate problems and what magnitude of impact these problems have throughout the system. Therefore, it is necessary to find a way to gather evidence of the problems in the most effective way possible, with the least detrimental impact possible on the treatment of patients.

The problems need to be stated in such a way as to clearly explain the effect the problems are having on the system. To expect the staff to know how to do this without training is unreasonable. Therefore, the best way to collect the statements that will contribute to the eventual UDE statements that will be used in the subsequent TOC analysis is for trained TOC practitioners to conduct short interviews with the staff on an individual basis.

Twenty- to thirty-minute interviews should be conducted in a safe, private environment and the interviewees should be assured that their contributing statements will not be attributed to them personally and that the final analysis will not include the names of the contributors.

Prior to the interviews, the participants should be told that their contribution is not forming a witch hunt, that it is not a process intended to place blame on them or their colleagues. This obstacle can be addressed effectively by

giving the interviewee a brief description of three of the basic assumptions of TOC.

1. **All systems are simple**, if understood correctly.
2. **There are no conflicts in reality**, just different perspectives of reality.
3. **People want to do good**; this is especially true in healthcare and it is often the system or people's perspective of the system that forces them to behave in ways that are counterintuitive.

By briefly explaining that these are the assumptions of the process in which they are participating and that their contribution will be confidential, most participants readily agree to participate. They intuitively know that the system should and can be improved, and when that is done, they will be able to provide better, quicker services to more patients. Furthermore, the participants also agree that in order to change the system they recognize their need to participate and support the proposed process.

The participants are then asked to tell the interviewers about the problems (UDEs) they experience in their work lives. These statements are noted by the interviewers.

At this stage of the analysis, time is the current constraint of the interviewees. Therefore, it is necessary for the TOC practitioners to subordinate the interview schedule to the needs of the facility.

It is also necessary to interview a range of staff from executives and physicians, to nurses, technicians, and administrative and service support staff. As well as capturing statements from the vertical structure, it is also necessary to collect statements across disciplines. Many of the interviewees will represent both aspects of a facility.

Building the Current Reality Tree [9](#) (CRT) of a Facility

The purpose of the CRT is to determine the core problem of a specific system, which in this case is a single facility. A facility-wide CRT offers a comprehensive and very detailed “snapshot in time,” clearly showing the interconnectedness of the problems that are being experienced by the staff and patients.

Because words are used as the primary source of “data,” a CRT easily incorporates details of behaviors, operational issues, policies, and protocols. Numbers can be included if they are needed to substantiate certain points, but the product of a CRT is written explanation of the existence of everyday problems and their source.

The process of building a CRT starts with writing the cause-and-effect logical relationship between closely related UDEs. Continuing to incorporate all UDEs in this way offers the readers a unique, revealing overview of their organization and a chance to recognize systemic patterns of behavior being exhibited by the staff, and understand why they exist.

Converting the Interviewees’ Statements into UDEs

Many of the statements collected during the interview process will be duplicates. These are easily collated and represented as a single UDE. Some statements appear to be standalone comments. Often, these take the form of a direct quote from a participant and, whenever possible, should not be generalized.

No statements should be dismissed at this stage, as they may be critical to the analysis, no matter how far-fetched they may appear. Some statements may even appear to describe a positive rather than a negative, but if the contributor considered it a negative, it should be included in the next step of the analysis to verify its orientation.

Constructing the CRT

Using rigorous cause-and-effect logic and the *Categories of Legitimate Reservation* (CLR),¹⁰ the UDEs are connected to reveal the core problem, which can then be expressed in the form of a Cloud to describe the core conflict ¹¹ of the facility.

During this process, it is necessary to have ongoing contact with a champion at the facility who is used to verify the logic used to construct the CRT, the core problem, and the core conflict.

Sphere of Influence

Constructed correctly, the CRT will identify both internal and external constraints. During the reporting process, and after the CRT, the core problem and its underlying core conflict have been verified, it is necessary to make staff at the facility aware of the need to plan to work within its current sphere of influence, the recognized bounded areas of activity over which the staff, including the executives, have the authority to make autonomous changes. A facility may be suffering from a legislative or corporate constraint that its staff has no current ability to influence. To try to do so at this stage will be a waste of effort and time that is needed to treat patients.

However, the ability to address corporate constraints will improve once the facility's executives are able to demonstrate that its ability to improve further is being blocked by corporate policies, by which time the head office will be keen to understand how the facility has been able to produce marked improvements in patient Throughput.

Most facility CRTs will expose many erroneous behavioral issues that are being driven by behaviors, policies, and procedures ¹² that eventually will need to be addressed. The danger at this stage of the process is that the staff will want

to address these issues in isolation—in effect, reverting back to addressing symptoms as opposed to the core problem.

How to Cause the Change

Training the Process Units

Once the core problem and its underlying conflict and their causal relationships to the numerous isolated UDEs have been identified and verified by the contacts (the champion and key staff) at the facility, it is time to begin to train the employees (managers, clinicians, and support personnel) to prepare them to overcome the facility's core problem.

The training needs to include an overview of TOC and how it addresses problems. The people who need to be trained are those working at the facility who will be needed to introduce the changes necessary to overcome the systemic conflict. Often this will require people from all levels of the facility to be trained as the CRT will clearly show the far-reaching effects of the deep-rooted core problem.

To this end, the training needs to offer trainees opportunities to work on existing problems through guided practice using the three basic behavioral TOC tools:¹³

1. The Evaporating Cloud¹⁴
2. The Negative Branch¹⁵
3. The Ambitious Target¹⁶—a derivative of the Prerequisite Tree¹⁷ developed by TOC for Education

The repeated use of these three tools will increase the ability of the staff to overcome many of the nonsystemic

and interpersonal problems they encounter during their working day.

The Process of Ongoing Improvement

Providing a Knowledge Base for Achieving the Goal Now

The Cloud

The Cloud will give them the critical thinking skills they need to:

- Make effective win-win decisions.
- Understand and facilitate their own and other people's understanding of situations.

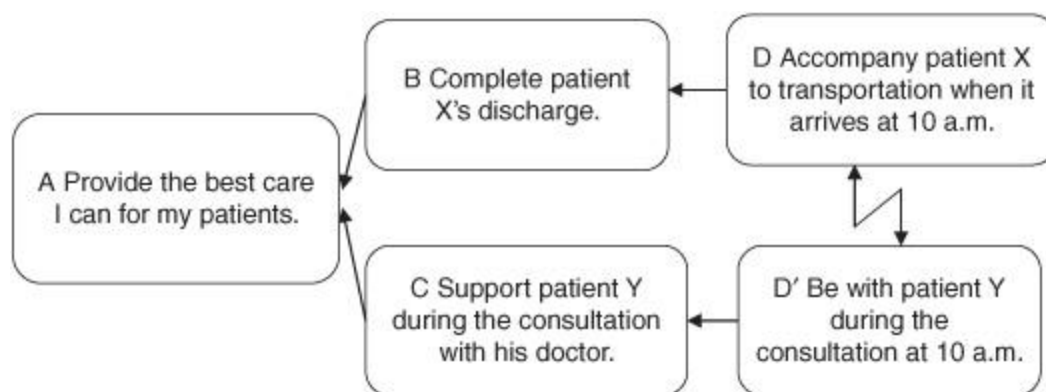


FIGURE 32-3 A nurse's dilemma.

- Resolve dilemmas and conflicts on many levels—personal, departmental, etc.
- Be receptive and willing participants in other people's or department's problems.

In [Fig. 32-3](#), we see an example of a typical problem in large-scale healthcare systems that many nurses experience.

When nurses are allocated patients, they are responsible for their care and are often expected to attend to many of their patients' needs. However, when the needs of two patients clash, they are often caught in the dilemma of who to take care of and are often forced to resolve this by delaying care for one patient in favor of another.

This cloud reads:

In order to [A] provide the best care I can for my patients, I need to [B] complete patient X's discharge, and in order to [B] complete patient X's discharge, I want to [D] accompany him to his transportation when it arrives at 10 a.m.

On the other hand, in order to [A] provide the best care I can for my patients, I need to [C] support patient Y during the consultation with his doctor, and in order to [C] support patient Y during the consultation with his doctor, I want to [D'] be with patient Y during the consultation at 10 a.m.

Obviously, the nurse cannot be in two places at once.

[Table 32-2](#) shows some of the assumptions between the nurses' *needs*, B and C, and *wants*, D and D'.

The Top of the Cloud B-D			
In order to [B] complete patient X's discharge I want to [D] accompany patient X to his transportation when it arrives at 10 a.m.	Because accompanying patients off the premises is part of my duties.	Because the transportation is booked.	Because there is no one else who can accompany him.
The Bottom of the Cloud C-D'			
In order to [C] prepare patient Y for the consultation with his doctor I want to [D'] attend to him at 10 a.m.	Because he is my responsibility.	Because I know his case better than anyone else.	Because he has no family or friends here to assure him.

TABLE 32-2 Assumptions for Attending to Patient X or Patient Y

By sharing this cloud, it became evident that this dilemma was a common occurrence at this particular facility. Most nurses said that they had resolved this cloud by challenging the assumption of “because the transportation is booked” between B and D by spending time rescheduling the booked transportation until after patient Y’s consultation with his doctor. That way they could fulfill all of their duties—meet all of their needs and take the best care they could of their two patients.

However, on hearing about this the patient transportation service was eager to share how this particular resolution affected them.

The Negative Branch

The Negative Branch will give them a predictive tool to:

- Provide a process by which proposed solutions can be effectively critiqued.
- Differentiate and address the weak parts of proposed plans, thereby removing the need to reject them completely and improving on the original idea.
- Act as a communication tool for needed buy-in.

In [Fig. 32-4](#), we see a simplified example of how the transportation service used the Negative Branch to communicate its perspective of the problem of rescheduling patient transport times on short notice.

When the nursing staff read this Negative Branch and realized why there had been recent price increases for the services, they challenged the transportation providers need to increase the number of crews (the point at which the NBR turned negative) and therefore their charges. When it was explained to the nursing staff that approximately 20 percent of all of the discharge bookings from that particular facility had to be rescheduled, the nurses began to realize that an occasional change request from each of them was costing the facility more than they had budgeted for.

The nurses revisited their dilemma and decided to see if it was possible to find an alternative solution to rescheduling transportation on the day of the patient's discharge, especially as patients were usually disappointed when their return home was delayed.

The Ambitious Target Tool

The Ambitious Target Tool will provide a basic sequencing tool to:

- Offer a means to investigate the reasoning behind proposed actions.
- Bring a greater understanding of the need to sequence and protect the time to complete critical tasks.
- Give the staff a basis with which to plan their personal contribution to large projects.

[Table 32-3](#) is a simplified version of the Ambitious Target Tool that the nurses decided to use to take a closer look at the activities that take place during the discharge process.

They decided to challenge the assumption from the cloud of “accompanying patients off of the premises is part of my duties.”

When the nurses shared this information with the transportation managers, they were able to suggest that their crews were already qualified to bring patients into the hospital, which should mean that they were qualified to take them out.

Once this had been verified by the Legal Department, the transport crews collected patients from their beds and the nurses were able to say goodbye to their charges at their bedsides and use the time gained to take care of their other patients.

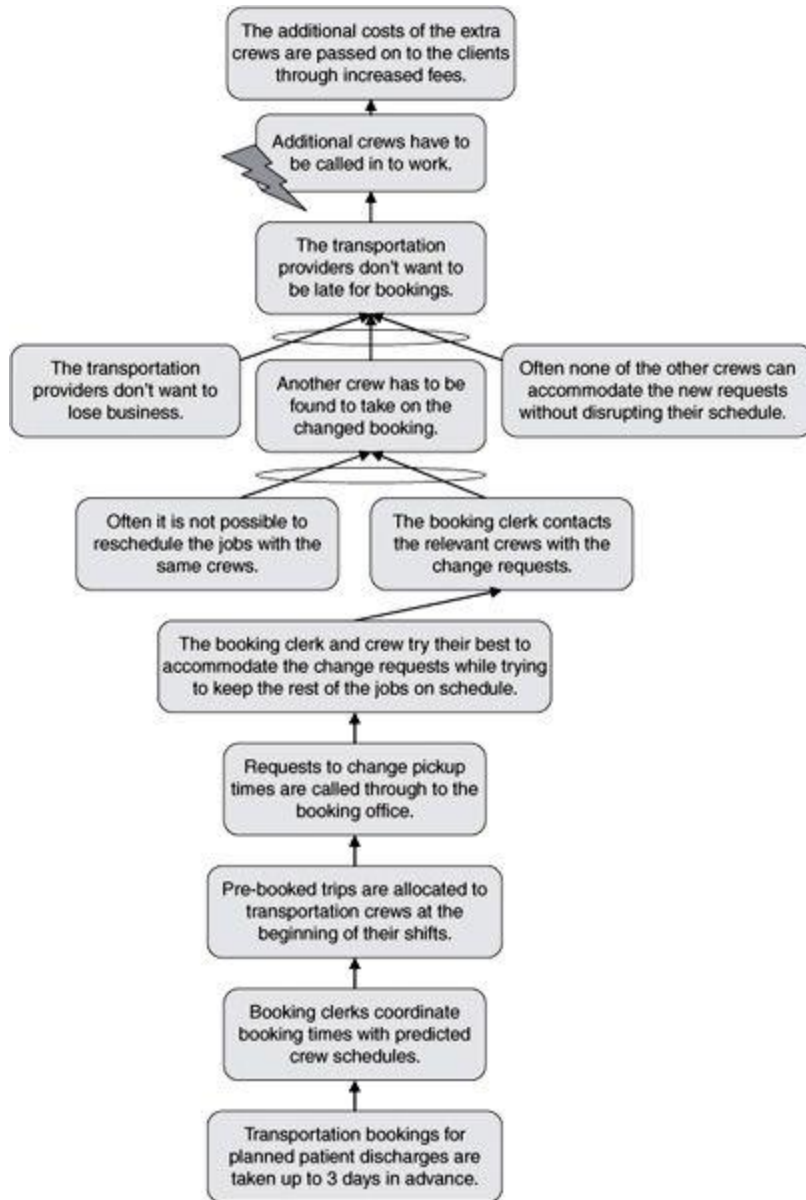


FIGURE 32-4 The problem of scheduling patient transport.

Be able to treat more patients, better, sooner, now and in the future		
Obstacle	Intermediate Objective	Action
Nurses feel duty bound to escort their patients off of the premises.	Nurses fulfill their responsibility to ensure that patients are safely escorted at all times.	Find someone else who is suitably qualified to escort their patients.
Patients have to be accompanied off of the premises to comply with hospital policies.	Patients are safely escorted off of the premises beyond which the hospital has no legal liability of care.	Find someone else who is suitably qualified to escort their patients off of the premises.
75% of the rescheduled patient discharge trips are due to nurses' time conflicts.	On the day of patient discharge transport needs do not cause time conflicts for nurses.	Nurses say their goodbyes to their patients before they leave the floor/ward leaving them plenty of time to attend to other duties.
Employing additional crews increases the overall costs of transportation services.	Keep the cost of patient transportation down.	Don't employ extra crews.

TABLE 32-3 PRT Results Challenge the Assumption from the Cloud of “Accompanying Patients Off of the Premises is Part of My Duties”

The overall results were:

- A reduction in the cost of patient discharge transport.
- Extra time for the nurses to take care of their patients.
- Far fewer delays in planned patient discharges due to transportation.
- Far fewer patients and relatives being disappointed by unnecessary delays in discharges.
- An improved working relationship between the nurses and transportation crews.
- Earlier and more predictable bed availability.

Providing the Knowledge Base for Achieving the Goal in the Future

Each TOC student needs to use these three tools a sufficient number of times to integrate them into their everyday thinking skills to become comfortable with their use and for

the tools to become each student's preferred tool of choice when they encounter problems.

Once this has been achieved, they will be ready to participate in the construction of the systemic plan that will address the core problem of the facility. The students will be ready to assist in the production of desirable effects (DE) [18](#)—the antithesis of the original UDEs used to construct the CRT—which will be used to build the lower, facility-specific levels of the *Strategy and Tactic Tree* (S&T). [19](#)

Addressing the New Core Problem

By using some of the problems included in the CRT as worked examples in the workshops, the students will have become very familiar with both the CRT and their own facility's core conflict.

In order to move the facility as a whole, it is necessary to produce a systemic S&T using this knowledge to populate the lower levels of the tree, the actions of which will address the systemic core conflict and move the facility into a position where the champion and the staff will be ready to address the facility's higher aims of the tree including:

- Recognizing the need to protect the time of the staff that will be used effectively to improve patient Throughput.
- How the facility will be able to identify and release latent capacity.
- How to support the staff to introduce productive behaviors.
- How the application of scientific thinking can be applied successfully to soft systems.

The facility specific S&T will also include, in the higher levels, the identification and incorporation within the facility of the knowledge of the higher level TOC applications that will be needed to bring about systemic Throughput improvements.

Five Focusing Steps

The *Five Focusing Steps*(5FS) (Goldratt 1990, [Chapter 11](#)) is a systematic five-step approach used to improve a system's ability continually to obtain goal units. The steps are as listed in the following:

1. IDENTIFY the system's constraints.
2. Decide how to EXPLOIT the system's constraints.
3. SUBORDINATE everything else to the above decision.
4. ELEVATE the system's constraints.
5. WARNING!!!! If in the previous steps a constraint has been broken, go back to Step 1, but do not allow INERTIA to cause a system's constraint.

Critical Chain Project Management (CCPM)[20](#)

CCPM is the TOC solution for planning, scheduling, and managing performance in a project environment. It is applied in two very different environments—single project environments and multi-project environments where resources are shared across several different projects concurrently.

TOC Synchronized Supply Chain Application[21](#)

The TOC Distribution/Replenishment solution is a pull distribution method that involves setting stock buffer sizes and then monitoring and replenishing inventory within a supply chain based on the actual consumption of the end user, rather than a forecast. Each link in the supply chain holds the maximum expected demand within the average replenishment time, factored by the level of unreliability in replenishment time. Each link generally receives what was used, although this amount is adjusted up or down when buffer management detects changes in the demand pattern.

Drum-Buffer Rope (DBR)²²

DBR is the TOC method for scheduling and managing sequential process steps.

Buffer Management (BM)²³

BM is the TOC method of identifying the current status of items with respect to arriving at the bottleneck and causes of lateness in items of arriving at the bottleneck. This tool is used to focus both expediting and local improvement efforts, which results in global improvement.

The TOC Thinking Processes (TP)²⁴

The TP is a set of logic tools that can be used independently or in combination to address the three questions in the change sequence, namely, What to change? What to change to? and How to cause the change?

Leaving a TOC Legacy

The aim of this program is for the TOC experts to leave each participating facility with the knowledge to be able to

maintain the process of ongoing improvement, through repeated application of the 5FS and the in-house knowledge and confidence to use the TOC applications until they are able to position and manage their constraint in such a way that it maximizes their ability to strive for the goal of treating more patients, better, sooner, now and in the future.

Summary

Contrary to common belief, the quality and cost of healthcare delivery is far more dependent on the *people* delivering the service than the infrastructures in which they operate. Excellent medical prevention and treatment can take place in the most basic of settings if the people delivering it are well trained, knowledgeable, and have access to the supplies they need. However, expensive and well-designed large-scale healthcare facilities, infrastructures, and buildings can fail in their purpose to deliver good quality, affordable, and timely care if the people working within it are hampered by the way the internal systems operate. By failing to meet the needs of their patients and their staff, these organizations can stagnate and lose the ability to make effective improvements.

All too often improvement projects in large-scale healthcare systems fail to yield the expected results. More often than not this is not due to a lack of efficacy on the part of the methodology used, or a lack of the intent by the people trying to improve matters, but rather a lack of understanding of the underlying issues that need to be addressed to unlock the stalemate generated by so many failed attempts to progress matters. In addition to breaking the “improvement stalemate,” there is the added obstacle of the day-to-day business of the hospitals and clinics, which

cannot and should not be interrupted. Unlike a production line, it is not possible to shut down a clinic for a refit if the demand for its services cannot be satisfied elsewhere. Healthcare is a continually traveling carousel of activity onto which improvement programs have to leap and be successful *without* disrupting the daily business of providing care.

In order to provide the improved levels of care their patients need, operational health-care improvement efforts need to be subordinate to the day jobs of caregivers. The people working in healthcare have to be able to integrate changes that will bring about real gains with a minimum of disruption to patients and services. However, even before changes are attempted, the people expected to implement them need to be able to voice any concerns they have and contribute their own expertise and experiences about any proposed changes in the processes they perform each day. All too often, the operational knowledge and intuition of the staff is not sought or offered. However, giving people the opportunity to participate in the planning of improvement projects is insufficient.

In any Emergency Room, a team of well-trained, experienced medical and nonmedical support staff can treat multiple patients with incredible speed, accuracy, and high quality of care. Charge the same team with improving patient Throughput management in the Emergency Room and they will likely suggest as many ways to improve matters as there are people in the discussion. Furthermore, if there are physicians present the number of suggestions will likely double as they attempt to consider the merits of their own opposing views! So, what is missing? Why is it so difficult to gain consensus and implement successful improvement initiatives in healthcare settings? There certainly is not a lack of methodologies, intelligence, or ability. Quite simply, it is due to a lack of a common language and processes to resolve issues in a way that will

bring all of the participants to agreement without having to compromise any of the important needs of the stakeholders.

By producing a factual, system-wide analysis of how the prevailing problems are affecting the system as a whole and how these interactions produce ripple effects throughout the system, it is easy for the staff to recognize why certain difficulties arise out of the interactions between departments, divisions, and personnel. With this level of analysis, it is a simple task, often for the first time ever in the life of a facility, to demonstrate the way internal systems, policies, and procedures have evolved and why some of them are outdated or inappropriate for current needs, forcing people to behave in ways that are often counterintuitive and sometimes badly. Furthermore, this analysis can begin to open new lines of communication and repair those that are failing or have broken down. The provision of this platform and mechanism at the outset of a TOC improvement program in a large-scale healthcare system provides a very powerful demonstration²⁵ of how the TOC tools provide a mechanism to begin to effectively address UDEs experienced by the staff.

With such a high dependency on the behavior of people, the initial core problems of individual facilities are highly unlikely to be operational issues, but rather they will be behavioral. Of course, operational issues will exist in every facility, but addressing the deepest problem of a lack of platform and mechanism by which negatives can be raised and effectively addressed will yield far greater benefits when the constraint becomes an operational issue.

In partnership with the system-wide CRT, training the people in the three basic TOC tools provides the staff with the mindset they need to be receptive, decisive, and willing to participate in the development of new solutions to longstanding problems. Practicing these tools on small everyday issues clears much of the “noise” out of the

system to reveal the “skeleton” of operational issues residing in the original CRT analysis that need to be addressed.

By cycling through the 5FS and training trainers to disseminate the knowledge of the three tools within the facility, it is possible to rapidly achieve exponential improvements in all of the necessary measures that are desirable in large-scale healthcare systems—Throughput, cost, quality, and waiting times to be able to treat more patients, better, sooner, now and in the future.

Proof of Concept

The author of this chapter was able to apply these principals to a large not-for-profit health-care system that was able to:

- Triple patient Throughput with:
- Only a 5 percent increase in resources
- A sustained increase in service quality to over 96 percent
- A sustained increase in patient satisfaction to over 96 percent
- And achieved:
- Third-year operating profit (margin) equal to first year revenue

The organization had no difficulty in recruiting clinical staff and establishing a waiting list of professionals ready to work for the organization. It continues to provide the margin needed to achieve its mission today.

References

Goldratt, E. M. 1990. *The Haystack Syndrome: Sifting Information out of the Data Ocean*. Crotonon-Hudson, NY: North River Press.

Goldratt, E. M. 1999. *Goldratt Satellite Program Session 6: Achieving Buy-in and Sales*. (Video series: 8 DVDs) Broadcast from Brummen, The Netherlands: Goldratt Satellite Program.

Sullivan, T. T., Reid, R. A., and Cartier, B. 2007. *TOCICO Dictionary*.

http://www.tocico.org/resource/resmgr/files-public/tocico_dictionary_first_edit.pdf

Wright, J. and King, R. 2006. *We All Fall down—Goldratt's Theory of Constraints for Healthcare Systems*. Great Barrington, MA: North River Press.

About the Author

After training in the field of engineering and working for many years as an award-winning business troubleshooter in the UK, **Julie Wright** found TOC while augmenting her business skills as a mature student. At the Avraham Goldratt Academy in New Haven, Connecticut, she was given the opportunity to combine her business experience with her passion for healthcare and committed to a long-term personal goal that would lead to the publication of *We All Fall Down—Goldratt's Theory of Constraints for Healthcare*, which she describes as the “What to Change” of the TOC improvement cycle for healthcare. After successfully implementing her findings in the UK, she is now working out of Dallas, Texas as the Director of Education for TOC-Healthcare Inc., introducing TOC to large-scale healthcare facilities in the United States and beyond.

As a long-time volunteer for TOC for Education, and like most other TOC practitioners, she claims to spend far too much time travelling and working in front of a computer and far too little time exploring the wonderful locations she visits for TOC events and work.

SECTION VIII

TOC in Complex Environments

[CHAPTER 33](#)

[Theory of Constraints in Complex Organizations](#)

[CHAPTER 34](#)

[Applications of Strategy and Tactics Trees in Organizations](#)

[CHAPTER 35](#)

[Complex Environments](#)

[CHAPTER 36](#)

[Combining Lean, Six Sigma, and the Theory of Constraints to Achieve Breakthrough Performance](#)

[CHAPTER 37](#)

[Using TOC in Complex Systems](#)

[CHAPTER 38](#)

[Theory of Constraints for Personal Productivity/Dilemmas](#)

Here, examples of TOC implementation and benefits for particularly complex environments, such as the large for-profit corporations, not-for-profit organizations, and other environments are discussed. The TOC Thinking Processes, The Strategy and Tactic Tree, TOC measurements, the Five Focusing Steps of TOC, and other TOC elements are brought to bear in real case examples showing how they work together as an integrated system of tools for sustainable improvement. Wide-ranging applications include

manufacturing, a church environment and how to improve personal productivity. Integrating TOC with Lean and Six Sigma; why and how to do it is also covered.

In a large complex corporation, how can the flow of ideas needed for development, production, sales and distribution of a new product be planned and tracked across organizational silos? How can executives at the top know that ideas are flowing as they should and that interorganization commitments are being met? How can they see issues coming before it is too late to recover? These topics are covered in this section. Then, how are Strategy and Tactic Trees used to frame the strategic direction of a company moving to make dramatic improvement in profit? Generic Strategy and Tactic trees are discussed with one examined in detail to show how the strategy is shaped and then used to knit the organization together with a unified focus on its strategic direction.

CHAPTER 33

Theory of Constraints in Complex Organizations

James R. Holt and Lynn H. Boyd

Overview

What makes an organization complex? What are the unique problems of complex organizations? How can the *Theory of Constraints* (TOC) help solve those problems? This chapter attempts to answer these questions. We start by providing a definition of complexity and then describe the core conflict of complex organizations, which we believe results from the need for both continual growth and organizational stability. One of the defining characteristics of complex organizations is that they have many independently measured units that are all trying to maximize local measures. The significant problem complex organizations face is coordination of these independent yet interdependent units. We make the assumption that the independent units and departments within complex organizations use *Drum-Buffer-Rope* (DBR) and *Critical Chain Project Management* (CCPM) to manage their internal processes, and we propose a key injection —“Everyone in the organization who has a significant impact on Throughput is measured by the same simple measure (that aligns all the actions of the organization with the goals of the organization)” —and show how it invalidates several of the assumptions underlying the core conflict. TOC provides measures for supply chains to achieve

coordination. We believe that TOC Supply Chain measures can be used in complex organizations to create an effective method of coordination between units and departments and to give senior management new insight and greater ability to manage independent units. These Supply Chain measures are discussed and examples of their application to complex organizations are provided. We also address additional injections related to “Conflict Resolution” and “Resource Allocation” along with a negative branch about “Leadership Certification.”

Definition of Complexity

There are many possible ways to define complexity. For this chapter on TOC in complex organizations, let us define *complexity* not according to size of the organization, nor to its technology, nor flow complexity, but rather *according to its TOC complexity*. With this criterion in mind, we have the following four levels of organization:

Copyright © 2010 by James R. Holt and Lynn H. Boyd.

- Simple
- Complicated
- Complex
- Chaotic

If the solution for the organization can be implemented with a single TOC tool (such as DBR or CCPM), we will call it a simple organization. If the solution for the organization involves the interrelationship of two or more TOC tools (such as DBR and TOC Replenishment seamlessly integrated into one package), then we will call it a complicated organization. If the organization has many independent yet

*interdependent elements each needing an individual TOC implementation responding to ever-changing product demands, we will call it a complex organization.*¹ This typically occurs when the organization has many independent business units, each with its own independent resource pools and individual profit/loss statements and yet the whole organization's effectiveness depends upon the successful contribution of many of the other business units. Complex organizations are characterized by multiple interactive constraints in a quickly changing environment aggravated by many near constraints operating within local optima guidelines without a clear, overriding schedule. A hospital is an example of a complex organization. Doctors, clinics, wards, laboratories, pharmacies, nurses, maintenance, housekeeping, imaging, record keeping and other functions, departments and units all try to perform at their best independently while also having to integrate many of their processes for effective organization performance. On top of the hospital's normal processing of patients, there are numerous improvement and new product development projects going on. Projects require communication and coordination between departments. Each department's capacity can be divided into capacity to meet current patient demand plus protective capacity. Protective capacity must be available on demand to support the constraint (current T) when unplanned demand occurs or Murphy strikes. However, at times when not confronted with these demands protective capacity is available for ideas projects, that is, future Throughput. Recognize, however, that if management is in the Cost World, when a resource is idle it is a perfect candidate for any cost-cutting initiatives.

An aircraft manufacturer is another example of a complex organization. Producing and assembling current models and producing spare parts for out-of-production models is a

complex task, but added to it are the demands of numerous and continual product and process improvement projects and the demands of developing new aircraft. In addition, these companies may compete in a number of markets including commercial, military, aerospace, and others. Maintaining these two major processes, one focused on current T and the other on future T, requiring many of the same resources, is extraordinarily complex.

Entities or systems that go beyond the definition of complex organizations, that is, *those that are quite unpredictable and lacking even a rudimentary flow structure (such as in sociology or politics)* we will refer to as *chaotic systems*. A chaotic system has many processes, policies, and procedures that frequently change and confound the solution space. While elements of the TOC proven solutions may help find a solution in chaotic systems, there is not a generic solution. Such systems require the TOC Thinking Processes (TP) to develop a solution.

In complex organizations, as in most organizations, the various departments or units each focus on maximizing the performance of their own limited resources. This aspect of the definition of complexity can apply to small organizations, such as an elementary school (each teacher trying to do the best for his or her classroom), or huge organizations like a country's armed services (and many other types of organizations in between). One could argue that complicated organizations without a single strong, integrating TOC solution often appear as complex organizations. As a result, the solution for complex organizations may be helpful in some complicated organizations.

Major Problems with Complex Organizations

Complex organizations suffer the same problems with on-time delivery, quality performance, sales, inventory management, and unstable workloads that have been discussed in earlier chapters of this handbook. Those issues, however, are exacerbated by the problems of connecting the different parts of the organization and motivating local managers to maximize global Throughput rather than local measures. The problem of multiple interactive constraints in a quickly changing environment is aggravated by many near constraints operating within local optima guidelines and without a clear, overriding schedule. This problem is similar to the significant shift from single project CCPM to multiple-project CCPM (as discussed in previous chapters). In single-project CCPM, all conflicts are removed in the planning stage and Buffer Management (BM) manages the task variability in execution. In multiple-project CCPM, eliminating every conflict is a poor solution that is inefficient, too long, expensive, temporary, and impractical.²

In complex organizations, the different product divisions, departments, resource silos, supporting plants, subcontracting entities, and other individual business units within the organization cannot all be synchronized for the same reasons that multiple projects cannot be synchronized in multiple-project CCPM. In addition, while the resources in one business unit are not always available to assist other business units, the effectiveness and efficiency of the overall organization are tied to the availability and allocation of resources and cooperation between all interested elements. In the economics literature, these problems are referred to as diseconomies of scale—problems of communication, coordination, and control that eventually overwhelm economies of scale as organizations grow larger and cause companies to become uncompetitive. Let's focus on those areas.

Undesirable Effects of Complex Organizations

Complex organizations frequently experience these typical *undesirable effects* (UDEs) to a varying degree:

1. Most of the experts are heavily overloaded.
2. Resources are not available when needed. Project managers hold on to experts, saving them for their next project.
3. Much work has to be done by less qualified experts.
4. There are too many needless delays.
5. Too often, the promised content is not achieved.
6. It is very difficult to respond quickly to every customer demand.
7. There is a frequent mismatch between what the customer wants and the resources available to perform the tasks.
8. Sometimes, our very expensive resources are left idle.
9. Promises are made without confidence in our ability to deliver.
10. The organization is not sure it will be able to deliver everything the customer wants.
11. Marketing frequently feels compelled to make promises that the company cannot keep.
12. It is difficult to replicate what appear to have been random successes.

13. Our reputation is tarnished.

The Core Conflict for Complex Organizations

The underlying conflict with complex organizations is the need for both growth and stability. [Figure 33-1](#) illustrates the goals of growth and stability drawn by Eliyahu M. Goldratt on many occasions and in many venues (Goldratt 1988b). These apparently conflicting needs lead to actions that force the organization in opposite directions.

Growth versus stability curves can be better communicated in an *Evaporating Cloud* (EC). [Figure 33-2](#) shows the objective, needs (necessary conditions), and wants to meet the objective and highlights a few sample assumptions that block complex organizations from achieving both growth and stability.

In order to have [A] a successful company, the company needs to have [B] continuous growth. In order to have [B] continuous growth, the company must [D] continually acquire additional capability. On the other hand, in order to have [A] a successful company, the company needs to have [C] stable operations. In order to have [C] stable operations, the company must [D'] avoid disruptions to the current capability. On one hand, the company must [D] continually acquire additional capability, while on the other hand, the company must [D'] avoid disruptions to the current capability. The company cannot do both simultaneously.

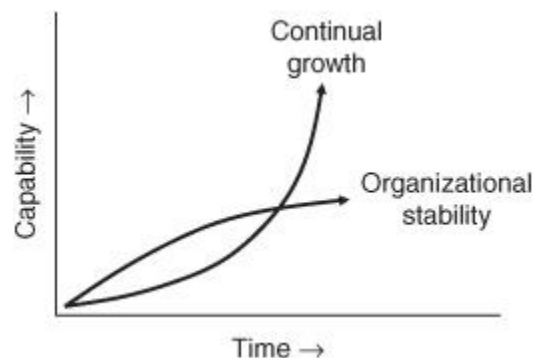


FIGURE 33-1 Growth versus Stability Curves. (Previously referred to as Red Curve—Green Curve and carried a different meaning. © E. M. Goldratt (1999) used by permission, all rights reserved.)

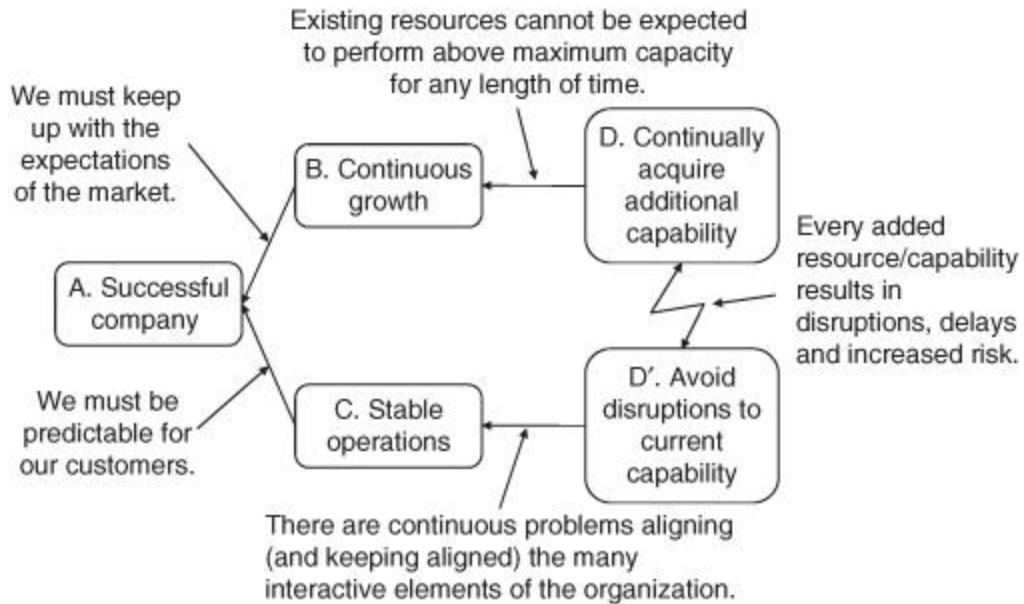


FIGURE 33-2 Core conflict.

The Direction of the Solution

The solution to a core conflict comes from examining and invalidating the assumptions behind the necessary logic.

What the Market Expects (A←B)

A [A] successful company must have [B] continuous growth because *the market expects it*. Complex organizations have external and internal customers. Few people are interested in a company that has declining growth or unstable performance. Publicly traded companies must maintain

continuous growth in value and profits to retain (avoid declining) stock prices. In addition, because the internal elements of the organization depend so much upon each other, the organization's success in one department depends upon improvements in other departments. For both of these reasons, the direction of the solution must support the ability to [B] continuously grow.

Adding Capabilities (B←D)

Attaining [B] continuous growth over time necessitates [D] continually acquiring additional capabilities or resources because *the existing resources cannot be expected to perform above maximum capacity for any length of time*. Internally, continuous growth causes continuous problems. While some parts of the organization can grow rather quickly, other parts of the organization cannot. Improvements in one area may be cheap and easy (and seem obvious); others may be expensive and time consuming (and not so obvious). Trying to keep the organization in balance (with ample protective but not excess capacity) demands the ability to grow everything at a rate that synchronizes the required contribution of each organizational element. The direction of the solution must address where and when to add additional resources to support effective [D] continually acquiring additional capability.

Predictable Response to Customers (A←C)

A [A] successful company must have [C] stable operations because *customers require predictable responses*. Unpredictable delivery reduces the value of the product offering and lowers market share. If we do not provide the delivery performance expected by our customers, they tend

to find someone else who can. However, what about internal customers? In complex organizations, other departments, offices, or functions also require predictability from each other. Failing to meet internal promises is probably more destabilizing for complex organizations than missing commitments to outside customers, even though the outside customer is not aware of it. Therefore, the direction of the solution must provide [C] stable operations.

Avoiding Disruptions (C←D')

Maintaining [C] stable operations by definition means [D'] avoiding disruptions to current capability because in complex organizations there are continuous problems aligning the capabilities of the many interactive elements of the organization. Complex organizations have many changing product mixes and varying workloads, which draw upon many interactive constraints. Even in the best of conditions, it is a terrible challenge for each part to live up to its commitments to other parts. Through no fault of its own, the overlapping demands from several critical and simultaneous endeavors can easily result in a department changing from having very little work to being heavily overloaded in just a few weeks. If there is no work, the expensive resources of the group seem excessive and costly. When there is too much work, there are often delays or quality problems. To many parts of the organization, it seems that just as one part gets in control, there is a disruption somewhere else that sends waves of work through the organization causing huge problems. For these reasons, it is critical that the direction of the solution must include a method to keep all parts of the organization aligned (in balance) to [D'] avoid disruptions to current capability.

Doing Both (D↔D')

We really have a dilemma when we must [D] continually acquire additional capacity yet at the same time we must [D'] avoid disruptions to current capability. This occurs because *it seems every added resource or new capacity disrupts, delays, and increases the risk of maintaining stable operations or supporting continuous growth.* A large part of complex organizations are individual business units that operate in their own best interest. They continually adjust their capacity through hiring and laying-off, building or shutting down, expanding or relocating. Acquiring highly technical professionals is a long lead-time problem, as is cutting back on expensive human resources. While these changes are somewhat disruptive to the individual business units, if there is an organization-wide requirement for continually adding capacity, the disruptions are magnified. Individual business units can actually compete against each other for a scarce resource pool. Units that try to reduce their local costs often cannot deliver to the changing demands from both inside and outside the organization, leading to delays and other problems. The direction of the solution must resolve this conflict in such a way that [D] continually acquiring additional capacity and [D'] avoiding disruptions to current capacity are not in conflict. Any added capacity must actually promote both stable operations and continuous growth.

Additional Understanding of Complex Organizations

Complex organizations continue to exist in part because they have good overall strategies. Without a reasonable strategy, the many challenges they face would quickly destroy the organization (or convert them to something less than complex). Strategies at the top of the organization are in many cases sufficient; however, as you go lower and

lower in the organization the interactions of the various organizational elements become much more complicated. Organizational elements, each trying to do their best, too often are at odds with one another. Conflicting goals between organizational elements at the lower levels not only block the lower elements from performing at their best, but also jeopardize the effective operation of elements above.

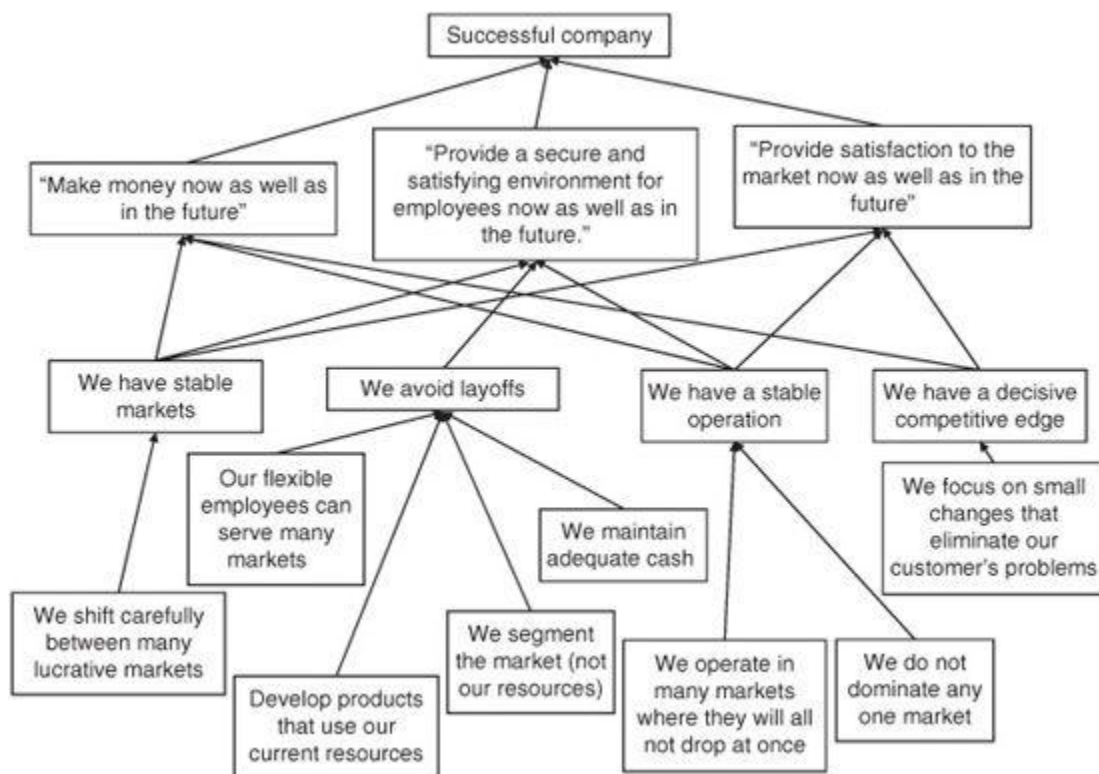


FIGURE 33-3 Generic strategy.

As an example, let's examine [Fig. 33-3](#), which is a very good generic strategic plan taken from Chapters 30 and 31 of *It's Not Luck*, by Eliyahu M. Goldratt (1994).

The generic strategy addresses the necessary conditions of the owners, the employees, and the customers. This is achieved by focusing on the customers' needs in such a way that competitors cannot quite duplicate, by choosing to service markets that do not all vary in the same direction at

the same time, by using employee resources in a flexible way, and by shifting between the lucrative markets of the time. Yet, even the best of strategies may not be implemented completely if lower down in the organization there are unresolved conflicts. This is shown in [Fig. 33-4](#), where lower down in the organizational structure there are conflicts that arise from different and competing measurement systems applied to different departments, organizational silos, independent business units, and employee reward systems.³

As we drop down through the layers of the organization to the tactics (objectives) of improving sales, accelerating projects, and improving distribution (a very small subset of the tactics involved in executing strategy in a complex organization), we see there are significant unresolved conflicts. The solutions to these specific individual conflicts⁴ have been addressed in previous chapters of this handbook. However, the problem is not quite so simple. What we see is that persistent unresolved conflicts at the bottom of the organization⁵ reflect back upward to higher levels of the organization to the point that there is conflict at all levels. [Figure 33-5](#) illustrates the resulting generic conflict that propagates upward to all levels (even to the CEO): In order to succeed, we must do those things that will allow us to succeed. However, in order for them (the other side) to succeed, we must not do the things we deem important for our success. These conflicts happen at every level. Unresolved conflicts from below propagate themselves upward to jeopardize or restrict the success of the company. For example, when plants or departments focus only on efficiency, they often restrict their focus to very few like products so they can gain the highest level of productivity. The department (or plant) then becomes very sensitive to any market downturn for their few products. A second example occurs when normal project fluctuations result in

unavoidable layoffs of people. However, when periodic layoffs are inevitable, we are not providing a secure and satisfying environment for the employees. Moreover, without dedicated employees, we seriously jeopardize the success of the company.

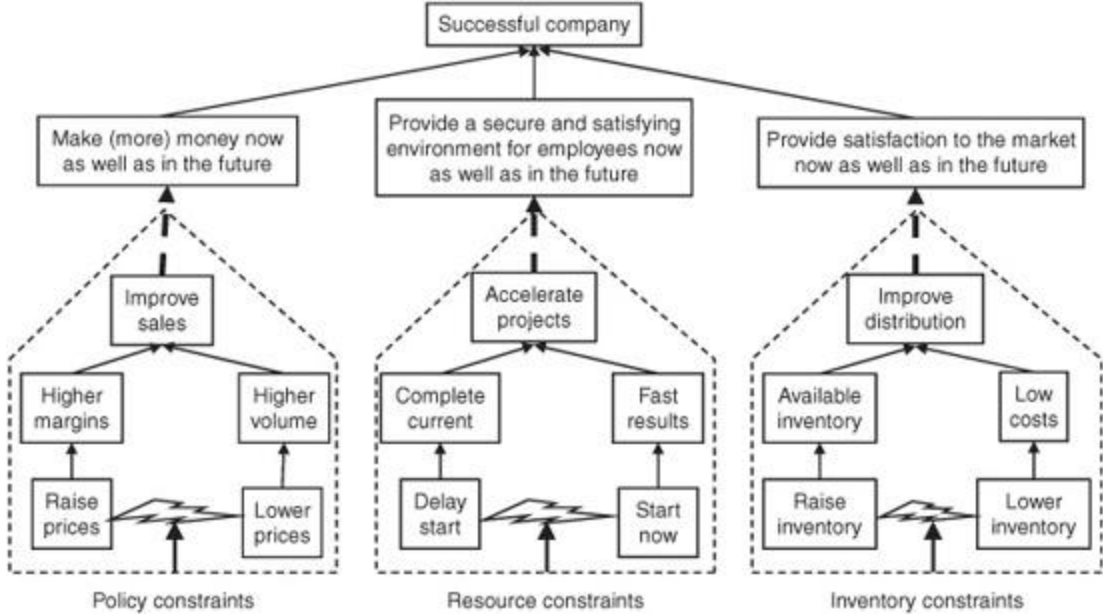


FIGURE 33-4 Conflicting tactics.

[33-2](#) are provided in [Table 33-1](#). Some potential individual injections are provided in [Table 33-2](#).

Looking at these assumptions and potential individual injections, it appears the complex organization is a complex supply chain (or maybe a supply mesh) with internal and external links. The problems of the complex organization mimic the supply chain and lead to the typical distrust between links and the over/under capacity problems experienced by the supply chain. However, these relationship problems are compounded by many more interlinkages than exist in a typical supply chain and the fact that transactions and requests between organizational units are not as clearly defined as market transactions and are difficult to prioritize with each unit's market transactions. While each element of the organization is trying to do its best, the problems continue. One cause for this is that many parts of the complex organization seem to have their own independent performance measurement systems. An example of this is sales measurements (such as keeping the sales funnel full) triggering over-commitment of development resources. Another cause is that different feedback or lag times and adjustment periods exist across the parts of the organization.

Arrow	Assumption
A ← B	We must keep up with the expectations of the market.
A ← C	We must be predictable for our customers.
B ← D	Existing resources cannot be expected to perform above maximum capacity for any length of time.
C ← D'	There are continuous problems aligning (and keeping aligned) the many interactive elements of the organization.
D ↔ D'	Every added resource/capability results in disruptions, delays, and increased risk.

TABLE 33-1 Assumptions of the Growth versus Stability Cloud

Arrow	Injection
A ← B	We can grow at a rate we choose.
A ← C	We always deliver to our promises.
B ← D	Our organization uses its resources effectively.
C ← D'	It is easy for offices, departments, resource pools, and plants to support each other.
D ↔ D'	We maintain an effective balance of stability and growth across the organization.

TABLE 33-2 Potential Injections of the Growth versus Stability Cloud

Breakthrough Injection

The breakthrough injection is selected by finding a single strategic injection that will satisfy all the individual injections and lead to all the needed desirable effects described to this point. A breakthrough injection is defined as everyone in the organization who has a significant impact on Throughput is measured by the same simple measure (that aligns all the actions of the organization with the goals of the organization).

Concepts in Organization Complexity

In order to understand this breakthrough injection and further define it, let's first review some important concepts associated with organizations as a whole and particularly associated with complex organizations. The four Supply Chain Flow Concepts developed by Henry Ford and Taiichi Ohno and interpreted by Eliyahu M. Goldratt (2009)⁶ are:

1. Improving flow (or equivalently lead time) is a primary objective of operations.

2. This primary objective should be translated into a practical mechanism that guides the operation when not to produce (prevents overproduction). Ford used space; Ohno used inventory.
3. Local efficiencies must be abolished.
4. A focusing process to balance flow must be in place. Ford used direct observation. Ohno used the gradual reduction in the number of containers and then gradual reduction of parts per container.

In complex organizations, the difficulty of achieving even the first Supply Chain Flow Concept is compounded by the existence of many interdependent specialists, departments, plants, offices, and resource pools. Keeping them synchronized is very difficult. Any optimization effort is short-lived. Before we talk about how the four concepts can be implemented, we need to be clear about two things—the types of activities performed within organizational units and the types of flows across unit boundaries.

Categories of Activities

Each resource or person will perform one or more of the following five types of activities:

- Day-to-day production—generally to meet current demand (Current T)
- Project activities—work on approved and scheduled projects (generally Future T)
- Idea development—work on developing ideas for future projects (Potential Future T)
- Support activities—to support the functioning of the unit or organization (very indirect relationship to T)

- Idle time—protective capacity (protects Current and Future T)

The mix of activities varies from unit to unit and, within units, from resource to resource. Most employees and machines in a manufacturing plant would have day-to-day activities related to current Throughput but might be called upon occasionally to contribute to projects. A quality control specialist in a manufacturing plant may spend 40 percent of her time on quality monitoring for current production and 40 percent on new product development projects to which her department has committed. Most of the resources of a development department might work entirely on either project activities related to approved projects or developing ideas for future projects, but would not have any day-to-day processing responsibilities related to current Throughput. See [Fig. 33-6](#) for some examples of time allocations for different resources.

In some units or departments, many resources will have a very direct impact on Throughput, while in other departments resources may have a very indirect effect on Throughput at best. We will use this relationship to Throughput later in this chapter to determine how departments should be measured, but for now [Table 33-3](#) shows how departments or units would be categorized by how directly they affect Throughput for the organization. We will see shortly how the categories shown in [Table 33-3](#) are useful in determining appropriate measures for each department or unit.

Flows in Complex Organizations

Now let us consider a simple organizational structure and the flows that cross the boundaries of different

organizational units or departments. The most common type of flow is that related to day-to-day processing or production to meet commitments for current Throughput to customers; for example, the flow of finished goods from Production to Distribution. However, there are many more interactions and complex flows related to projects. [Figure 33-7](#) illustrates that many people from all parts of the company contribute ideas to the development of a new product. For many such projects, there is a significant amount of interdepartmental discussion and flow of ideas before the project is formally approved. Once the project is approved, there is more interdepartmental flow associated with the activities that are part of the approved project. In [Fig. 33-7](#), a block labeled “Ideas” can refer to either an exchange between departments in the pre-approval stage or a department’s agreed upon obligation to do something to help deliver the final product of “Ideas.” For example, the Design department might have a preliminary design and has solicited “Ideas” on manufacturability from the Production department, essentially asking, “Can you manufacture this with existing resources?” Production, in turn, confirms that it can or provides necessary information to Design to allow the preliminary new product design to be modified. The two-headed arrow between Design and Production represents this interaction. The Sales department contributes comments from their marketing studies. The Service department contributes improvements learned from past products. Distribution suggests ideas for better packing and delivery. Design wants to incorporate the newest and best into the new product. Development includes its own ideas on how to deliver the best ideas they can to benefit the whole organization.

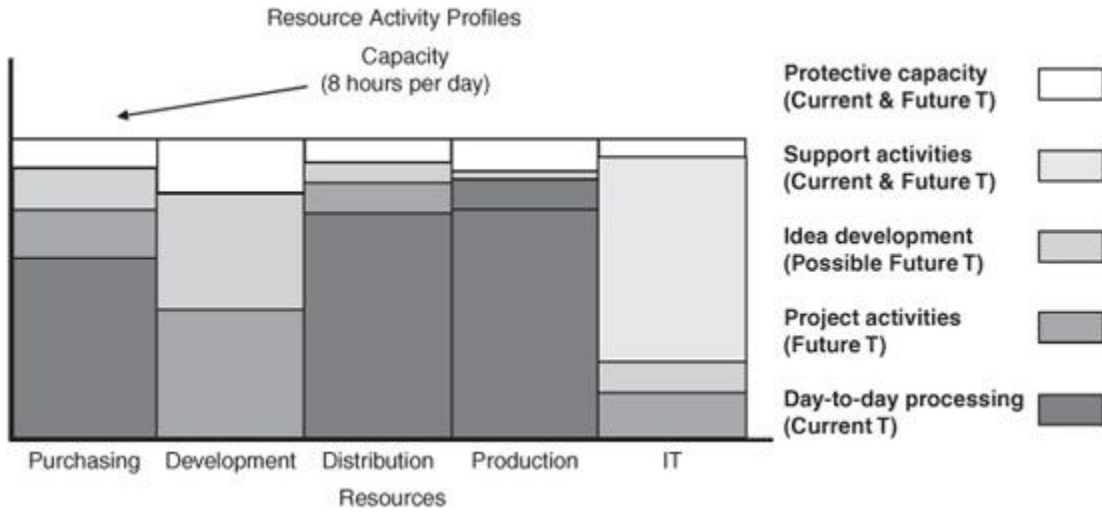


FIGURE 33-6 Resource activity profiles.

	Degree of Impact on Throughput			Units That Control Their Own Demand
	Primary	Secondary	Tertiary	
Examples	Sales	Human Resources	Financial Reporting	Process Improvement
	Production	Purchasing	Building Maintenance	Cost Reduction
	Distribution	IT	Legal	Internal Audit
	Engineering	Engineering		

TABLE 33-3 Levels of impact on Throughput.

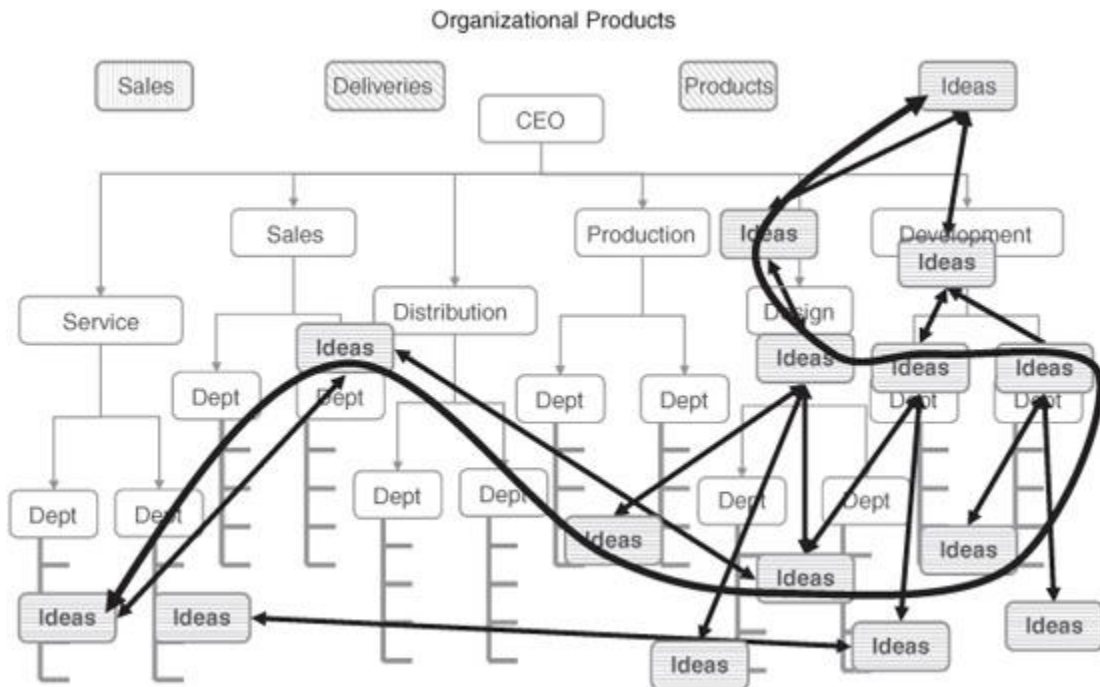


FIGURE 33-7 Flow of ideas.

[Figure 33-7](#) illustrates just a few of the flows across departmental boundaries necessary to develop an idea that may end up as a new product. When we add the flows for all other projects, both in the developmental stage and scheduled, and day-to-day production flows, the result definitely looks chaotic, as shown in [Fig. 33-8](#).

Looking at the complex flows in an organization in this way makes it clear why the management of such organizations is so difficult.

Flow Control with Critical Chain

In the earlier chapters of this handbook, there are solutions for different types of flow patterns: project flows, production flows, distribution, and sales. Using TOC Tools, we can tame this wild set of processes individually, yet they still need to be knitted together for an effective complex organization.

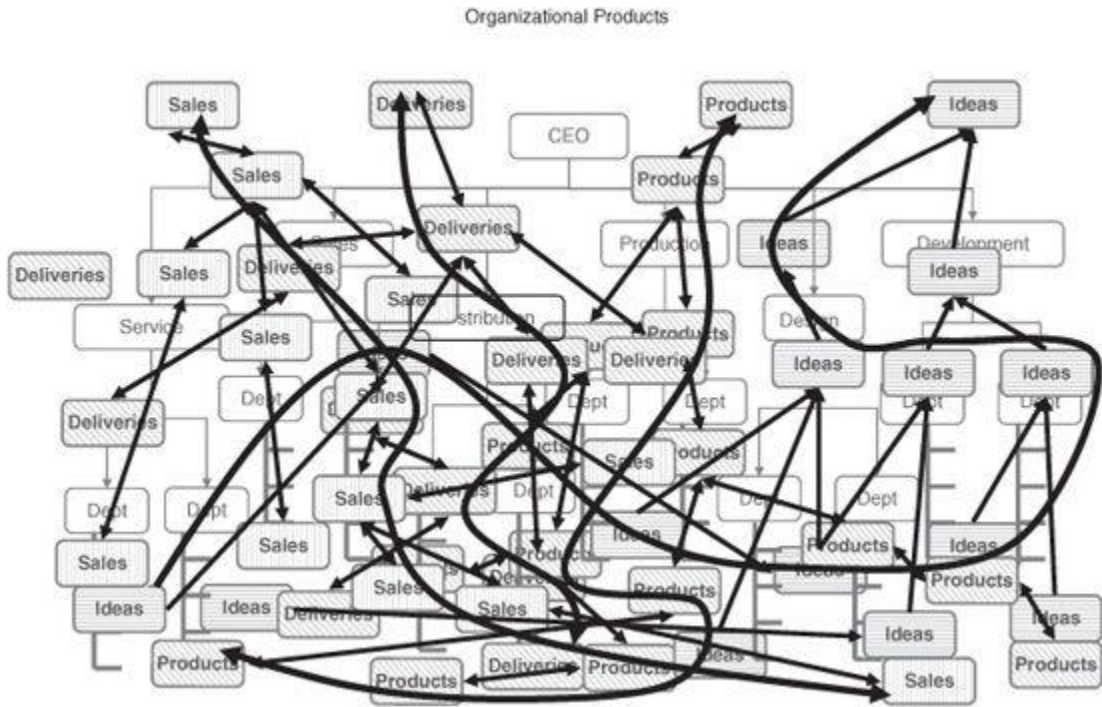


FIGURE 33-8 Flow of all products.

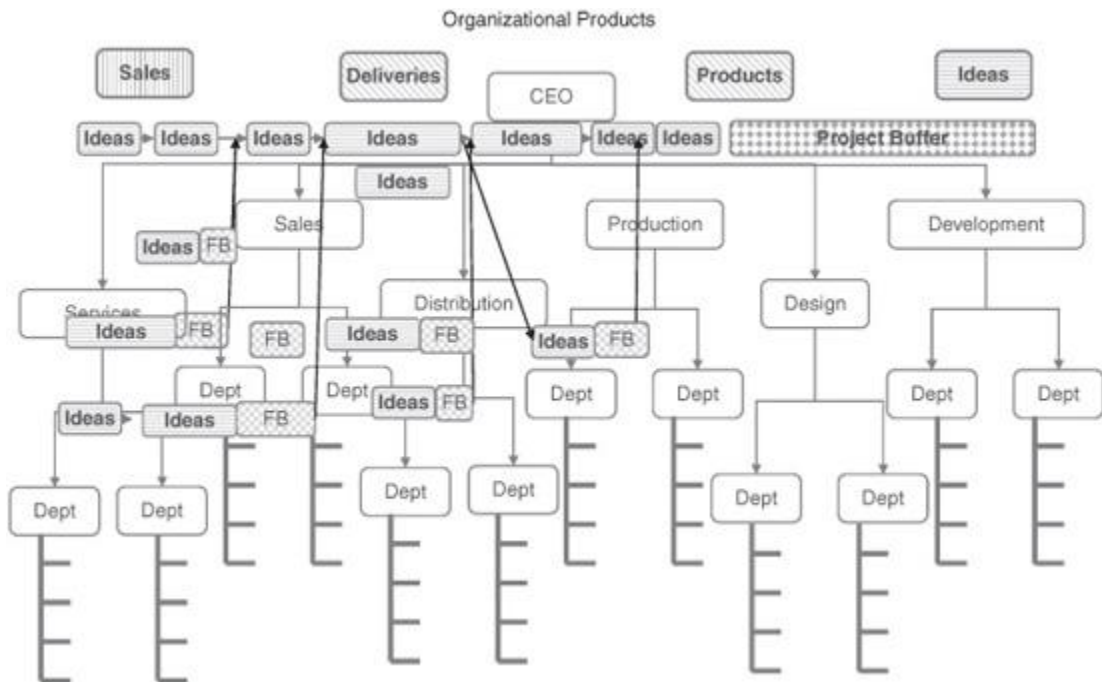


FIGURE 33-9 Buffered project for ideas.

The first of the four Supply Chain Flow Concepts (Goldratt, 2009) is improving the flow. If we take the flow of Ideas as shown in [Fig. 33-7](#), we can reconstruct the flow using the CCPM approach. Changing the length of each task to represent the expected task duration with an aggressive schedule shows the critical chain and the non-critical feeding activities.

A project buffer is included and feeder buffers are inserted. [Figure 33-9](#) shows such a plan.⁷

Adding similar solutions for Sales, Production, and Distribution using other TOC recommended processes (aggressive schedules with strategically placed buffers of each type) seems to offer an orderly set of templates for managing the organization. These different templates are shown in [Fig. 33-10](#).

Having an aggressive yet carefully buffered flow pattern for each process and project is not sufficient. If the system is allowed to run without control, every process tries to do its individual best until the system turns chaotic again as in [Fig. 33-11](#).

With so many things going on at once, even when the flows are well defined, the system is unwieldy. There must be some order. The second Supply Chain Flow Concept (Goldratt, 2009) is to insert a practical mechanism that tells the operation when not to produce in order to prevent overproduction. Ordering the flows in logical sequence driven by an overall customer demand such as in [Fig. 33-12](#) can reduce the chaos, speed reliable flow, and more effectively use internal resources. Ordering the flows in this manner means choking the release of work in each area in such a way that work begins only when the system is ready for it. This is done by pipelining projects in multi-project critical chain and using the rope mechanism in DBR or S-DBR for day-to-day production flows. In both of these cases

(projects and production), buffers are used to manage the flows.

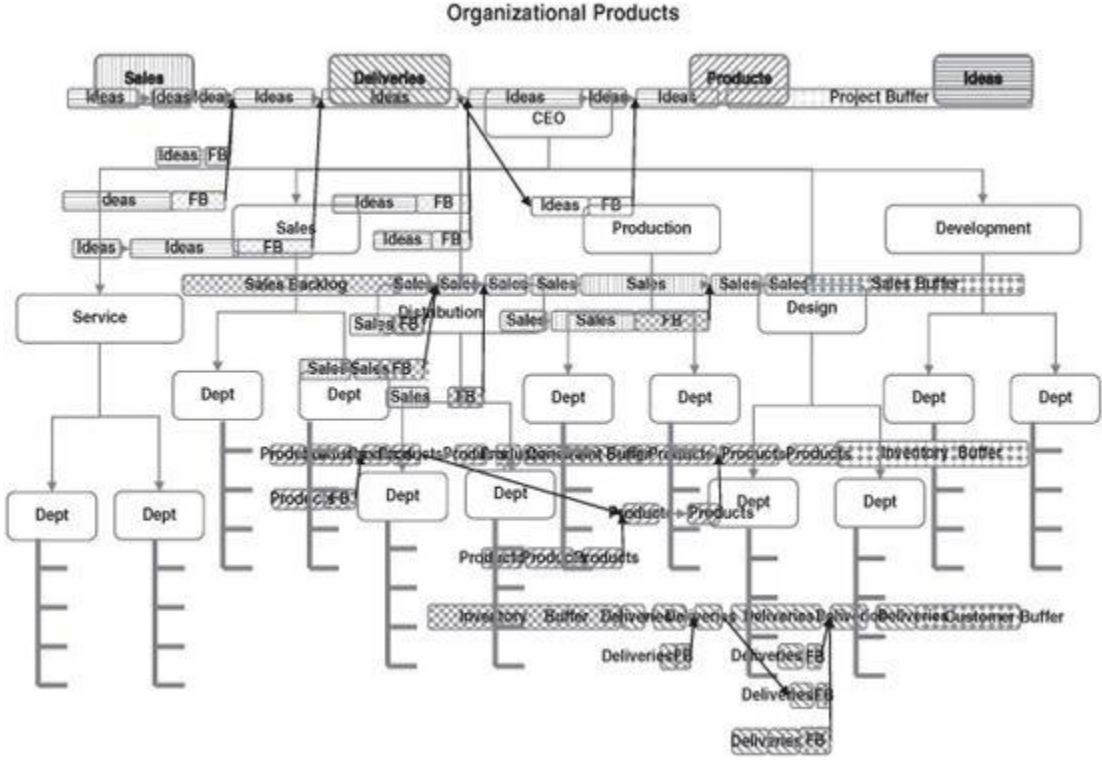


FIGURE 33-10 Buffered processes.

Multiple Project Chaos

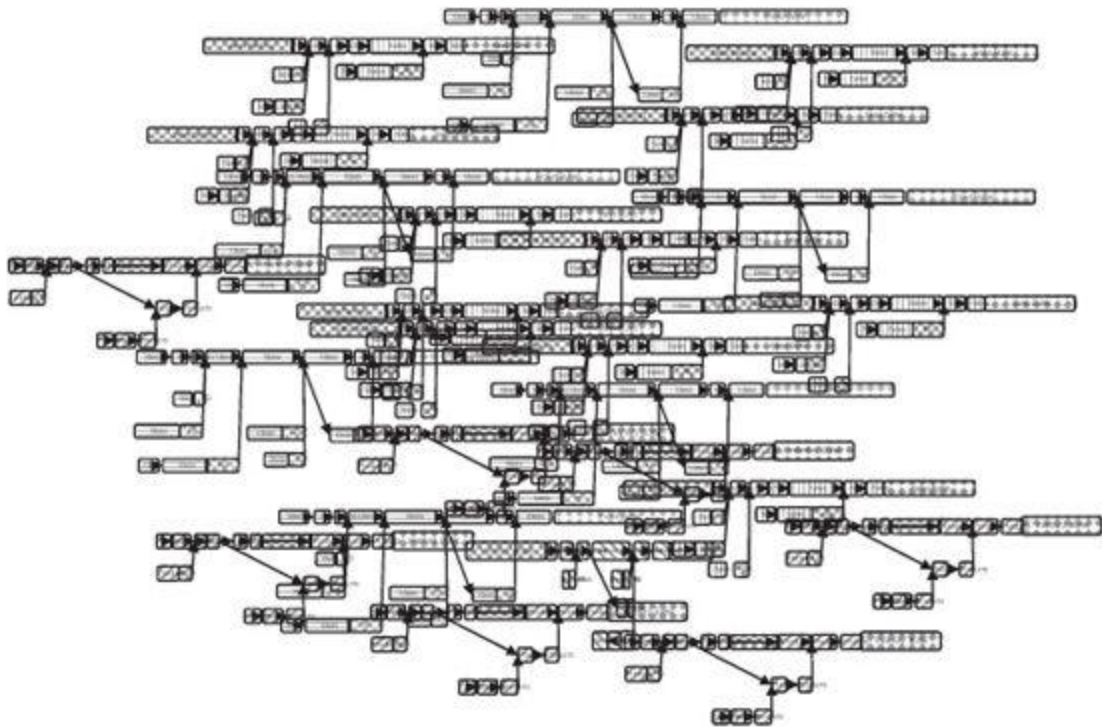


FIGURE 33-11 Multi-project chaos.

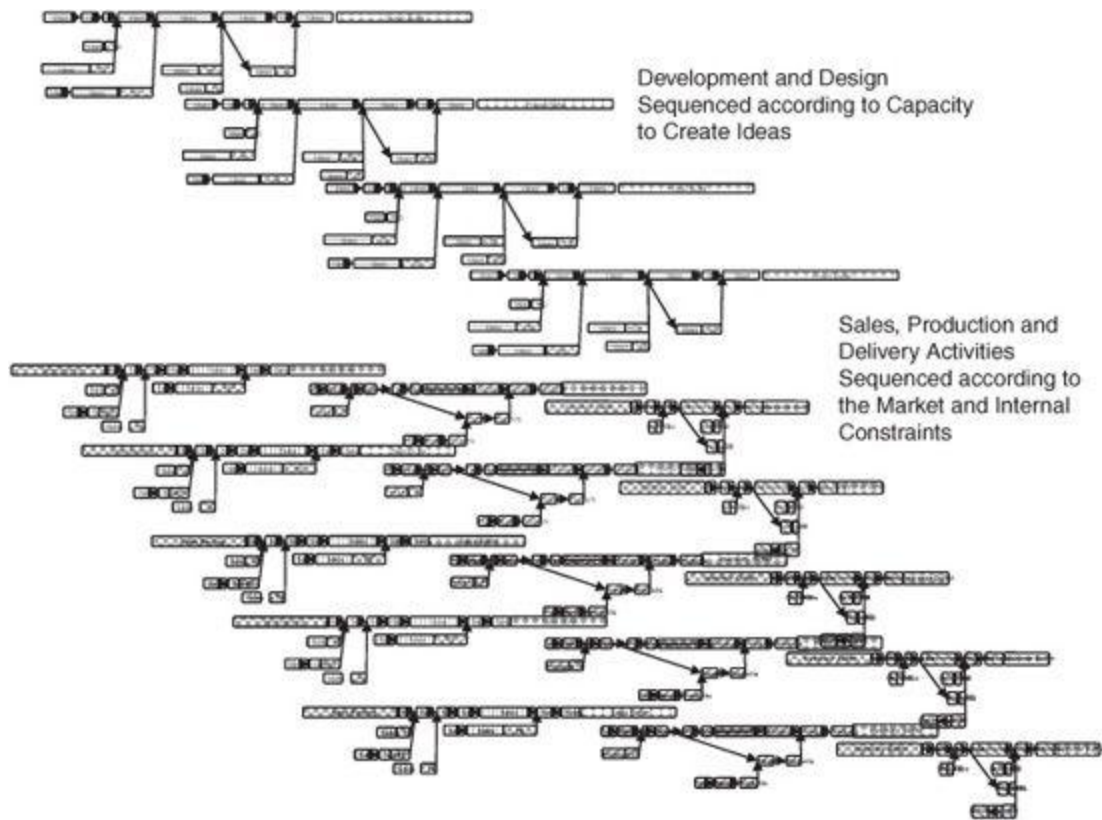


FIGURE 33-12 Sequenced process management.

Sequenced process management in [Figure 33-12](#) connects the flow patterns so that everything moves at the fastest capability of the organization but not too fast or too much. There are different flows, such as the development flow, that run parallel to the production flow. These different flows have cross-flow connectors needed for the different groups to support each other. This process flow arrangement also implements the third Supply Chain Flow Concept (Goldratt, 2009): Local efficiencies are abolished. Work is started and processed at the rate of the systemic constraint, not at the rate of local, nonconstraint resources. There will be some protective capacity in most departments, and the overall flow of the organization will be maximized.

A Breakthrough Injection

With this background, we are ready to revisit the breakthrough injection: Everyone in the organization who has a significant impact on Throughput is measured by the same simple measure (that aligns all the actions of the organization with the goals of the organization). With all the interactions of so many different departments and units, each with its own goals, objectives, and measurements, how can we imagine a single injection, let alone call it a breakthrough injection?

The Definition of the Common Simple Measure

Having buffers for production flows and CCPM schedules for ideas projects in each unit provides the basis for measurement. We want to measure the extent to which a unit is not doing what it is supposed to be doing. We could measure the “not done” things several ways. We could count the number of things not done. We could estimate the value to the organization of the things not done. Moreover, we could measure the duration of time until the error was remediated. Each one of these measures is insufficient. If we only count the number of errors, we could reduce many tiny errors and not really deliver value to the organization. Making an error with a large negative impact on the organization is bad. However, if it only lasts a day or so, it is not nearly as bad as a lesser error that lasts months. Reducing the average time to remove an error may be helpful, but that focuses improvement efforts on the minor, easy-to-fix errors and leaves the big errors unresolved. What we really need is a composite measure that considers both of these factors. Here we can benefit from one of the measures used in TOC Supply Chain Management: Throughput Dollar Days (TDD).⁸ TDD takes into account both the amount of Throughput that is delayed and the

length of time it is delayed. *TDD is the Throughput Value (the contribution to the organization represented by the final sale less any truly variable costs), which is assigned to product or process flow, times the number of days late added together for all late tasks (lateness is defined as not delivering the required quantity or quality on the mutually agreed upon date). TDD is a measure of reliability. It measures a department's or independent business unit's delivery to promise for both production flows and project activities.*

Before we discuss an example, we should be clear that we are interested in providing senior management with information that is useful for managing the whole organization. Within each separately measured unit of the organization, we expect managers to use TOC solutions (primarily DBR and CCPM) to manage their resources to meet their commitments. We are proposing here that TDD only be assigned at the point that either a production flow or project moves from one separately measured unit of the organization to another, that is, at the handoffs between separately measured units. We are using TDD to measure unit performance, not individual task performance within units.

It is also important to state that we are not suggesting that a milestone be established for every activity in a project. We are trying to provide senior management with information about separately measured units in complex organizations. A major problem in organizations that have significant project work is understanding the relationship between capacity and demand. In units that have both production and project work, most resources will generally work on one or the other. Occasionally a resource will be required to work on both production flows within the unit and project activities for other parts of the organization. For example, a test engineer may have day-to-day quality control responsibilities for a unit's production and be

required to regularly complete activities that are part of projects managed by other units of the organization. If the test engineering resource is a non-constraint, BM within the unit will ensure that both process flows and project activities are completed as required. If the test engineering resource is a constraint, use of DBR within the unit should prevent commitments either to an external customer or to another organizational unit that cannot be met. However, unit managers frequently have difficulty assessing whether they have resource constraints, and consequently, frequently make commitments that cannot be met. This is particularly a problem when project activities are a source of demand. In these situations, TDD will provide useful information to senior management as to where attention should be directed.

Using TDD to assign lateness to different departments and independent business units is similar to the TOC Replenishment solution for Supply Chains. It is not the same as normal CCPM. In CCPM, the project manager uses BM to take care of task variations. The project manager can allocate resources as needed to deal with buffer penetrations. However, in complex organizations, the resources are not generally available for reallocation. In complex organizations, the workflow is secured by promises that are critical to synchronization.

Most complex organizations make extensive use of milestones. Milestones are not so much a measure of performance as they are markers to indicate a project or process has progressed to a certain point. Milestones are not very good management tools because there are too few of them and they are too far apart. In addition, milestones are lagging indicators: They do not tell you what is coming in time to make corrections. They only tell you when it is too late to do anything about it. A missed milestone usually means a serious problem has occurred. Missing a milestone puts tremendous peer pressure (if not management penalty)

on the errant party. To avoid this peer pressure, groups will often inflate their delivery times and create all the problems discussed in Chapters 3, 4, and 5 on CCPM.

In contrast to milestones, TDD is a forward-looking indicator of what is going on in the organization. Units will begin incurring TDD before commitments to customers are missed because production flows and project activities that result in TDD are still buffered by shipping and project buffers. In addition, TDD is reported often. The periodic TDD by unit shows which units are having the hardest time delivering for a period of time. The TDD accumulating in a project or production flow can be seen early, well before it is too late to respond. When monitored over time, TDD (like the causes of buffer penetration) highlight where the organization needs to focus its management attention and improvement efforts.

With TDD, senior management will have a method to determine which departments or units need attention. Units will incur TDD when they commit to deliver things that they cannot deliver. A unit that incurs TDD will not automatically receive additional resources. The first question senior management should ask is why the unit manager has committed to deliver things that the unit does not have the capacity to deliver. While people figure out over time how to game almost any performance measure, the fact that TDD should be decreasing over time as unit managers become more proficient at DBR and CCPM means that units will not automatically be receiving additional resources just because they are incurring TDD.

Using TDD: An Example

Let's use an approved and scheduled project as an example. Many people and units are involved in projects but for simplicity, only three groups are shown in [Fig. 33-13](#). The

Service department is responsible for the first element of the non-constraint feeding chain. The Distribution department has two contributions, one on the subchain and one on the critical chain. The Production department has one contribution on the critical chain. For this example, assume the overall project is approximately 40 days on the critical chain with 20 days of project buffer. If the Throughput value of the final project “Ideas” is \$10, we can easily evaluate the subordination of the Service department, the Distribution department, and the Production department to the overall process of producing Ideas.

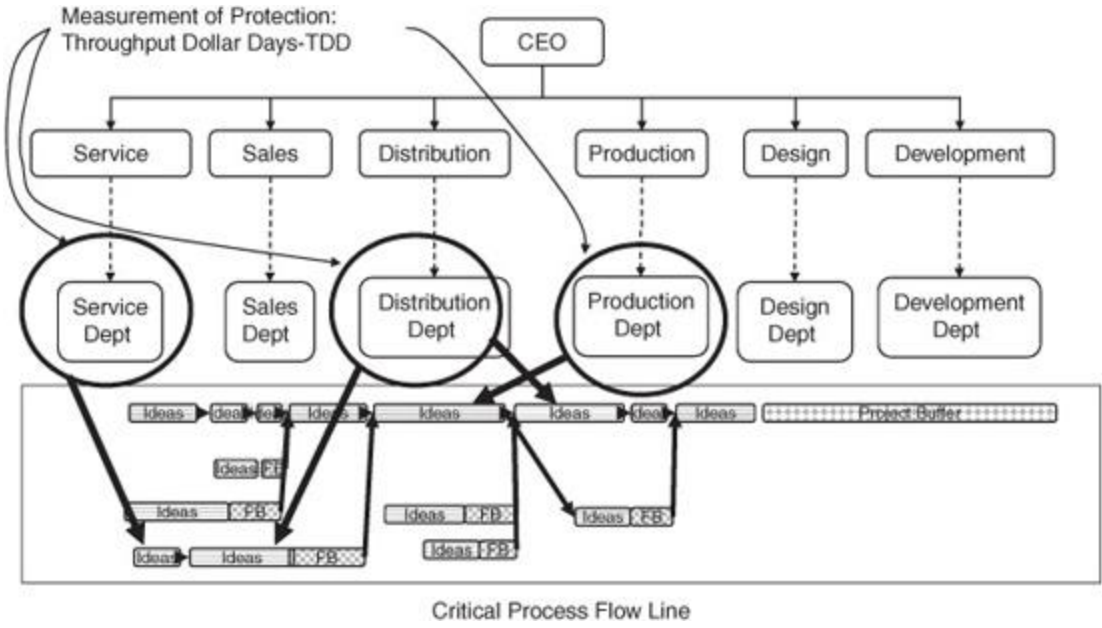


FIGURE 33-13 Distribution’s obligation to producing ideas.

Completing a project requires participation of many independent groups who agreed to delivery at a certain time or to respond within a certain period according to the plan. A unit or department may perform one activity or several consecutive activities on a project. The overall owner of the project illustrated in [Fig. 33-13](#) is the Development department. Because it is the project owner, the Development department has a project buffer that

protects delivery of the project to the final customer from any late delivery by the individual groups.

Let us assume the Service department agreed to deliver its part of the Ideas project on day five to the Distribution department, but actually delivered on day eight, three days late on a project valued at \$10 Throughput Value. The Service department is assessed three days \times \$10 or 30 TDD. Now, the Distribution department works on its task and completes it in one day longer than they planned. The Distribution department is assessed one day \times \$10 or 10 TDD. Both Service and Distribution departments have been assessed TDD. In thinking of TDD, we can say that each promised delivery is a commitment. When a promise is not met, that commitment is missed and the potential impact on the Throughput of the organization can be defined. For the project, the total TDD for this feeding chain is 40 TDD. TDD are additive. While TDD indicates a later delivery than promised, the normal TOC buffering process (the strategically located buffers used in DBR, CCPM, and Replenishment) protects against late delivery of the final product.

Next, we consider the Production department. Let's assume that additional actions by another department prior to the Production department have delayed the start of Production's task such that production is 10 days late to start. The department causing this delay would be charged 100 TDD ($\$10 \times 10$ days late). Production then takes two days longer than the response time they planned to complete their work. This additional two days taken by Production to deliver adds 20 TDD to Production and 120 TDD charged to the project.⁹

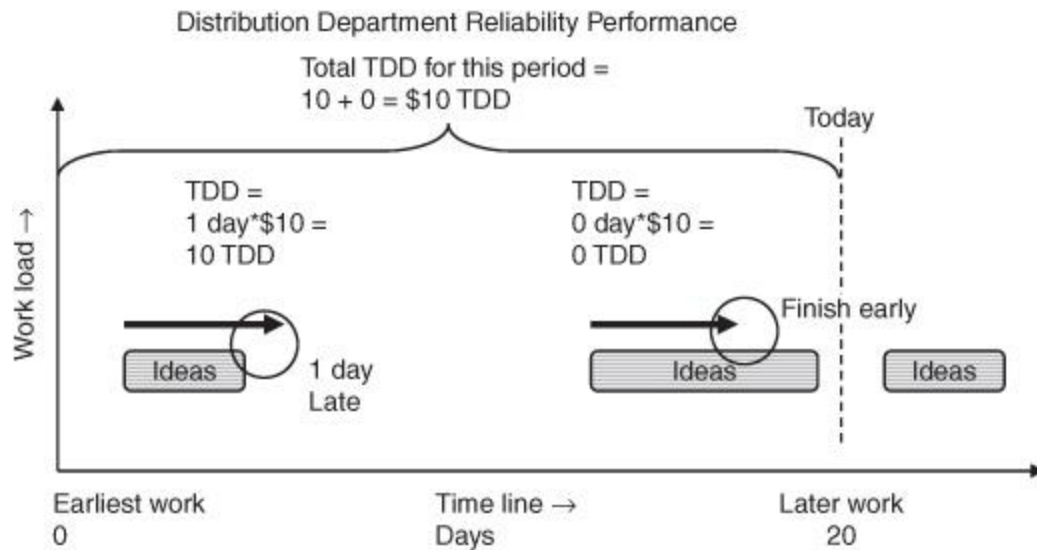


FIGURE 33-14 Distribution’s TDD performance.

The next critical chain task is performed by the Distribution department. By good chance, the Distribution department is able to complete four days faster than their commitment. The Distribution department is assessed zero TDD. The early completion by Distribution does not affect the TDD assessment for the department or the TDD assessed to the project. What does happen in reality is the early completion by the Distribution department helps recover some of the project buffer consumption, which is reduced from 12 days to 8 days.¹⁰

The TDD measures assessed to units along the critical chain paths are additive. They reflect the capability of the system (an accumulation of many individual groups) to deliver as promised. While TDD encourages independent yet interrelated groups to deliver to promise (to be reliable), it also gives management a much-needed measure to determine the capability of the overall system. Over time, TDD levels and trends can be used as an indicator of which groups need improvement, assistance, or elevation.

A Closer Look at the Distribution Department

Let us look at an example of how this works for the Distribution department. [Figure 33-14](#) shows the Distribution department's TDD over the most recent period of time (in this example, the period is 20 days) where the two tasks for Ideas (discussed earlier) were performed. The arrow above the task shows when the Distribution department actually worked on the project to deliver Ideas. Over this period, the first task for Ideas was one day late, incurring a 10 TDD assessment. The second task for Ideas was four days early and assessed zero TDD. The sum total for the period (just for Ideas) was 10 TDD.

There is a difference between commitment dates for separately measured units and task times for individuals. When individuals estimate their own task time, you cannot punish them for being late or reward them for being early. Either one (reward or punishment) motivates the individual to extend future estimates of task time. However, when TDD is used to measure the reliability of a unit's commitments to deliver, both internally and to customers, we expect the unit to use strategically placed buffers to ensure that the commitments can be met (see earlier chapters of this handbook, which address DBR, CCPM, and Replenishment). TDD is not a penalty but an indication to senior management that a unit is missing its commitments. TDD indicates how effectively the unit uses DBR and CCPM to manage its internal operations. If unit management has a good understanding of the unit's capacity and demand and effectively uses DBR to release work into the unit at the rate the constraint can process it, then TDD for the unit should be very low. TDD will be incurred, however, if management does not have a good understanding of capacity and demand and too much work is released into the system. The purpose of measuring TDD is not to punish or reward unit management but to tell upper management where attention

is needed. In addition, we do not reward a unit for recovering TDD on a project managed by another unit. Doing so would encourage units to overstate the number of flow days required or to increase unit resources without real need. Units with the highest TDD should not be punished. They should be studied and helped to see if improvements can be made. Tracking and managing TDD over time at the system level gives the signals needed by top management about the management of capacity at the unit level and can guide the growth process while maintaining stability.

Of course, the Distribution department contributes to more than just Ideas projects. Their predominant duty is for the distribution function itself, but the Distribution department also has assignments to support Sales and Production. [Figure 33-15](#) shows several segments of the Distribution department's workload over time. We see that the different products that they support cause some periods of heavy workload and other periods of less workload.

During the first 20-day period shown in [Fig. 33-15](#), the Distribution department had some trouble. They were one day late on the Sales task (at the top), one day late on the Ideas task (near the middle), and one day late on the Delivery task (the bottom). In total, over that period, there was a \$20 Sales task delayed for one day, a \$5 Delivery task delayed one day, and a one-day delay for the \$10 Ideas task; their TDD for the period was 35 TDD. During the second period, the Distribution department improved their primary performance measure. They had no late deliveries for zero TDD. We note that they were able to start some of this work early (they shifted the schedule) to take advantage of time when they were not so overloaded.

Units to Which TDD Applies: Degree of Impact on Throughput

TDD is clearly applicable to some units of complex organizations, such as production, sales, distribution, and engineering. These units have clear commitments to either outside customers or to other organizational units that directly contribute to Throughput. In addition, there is another category of units that can delay Throughput even though they do not make commitments to customers themselves. For example, the Human Resources department can affect Throughput if necessary employees are not hired and trained when needed. Purchasing can delay Throughput by contracting with unreliable or poor quality suppliers. IT can delay Throughput if it fails to deliver a critical application to Production on time. TDD should be recorded for these units as well.

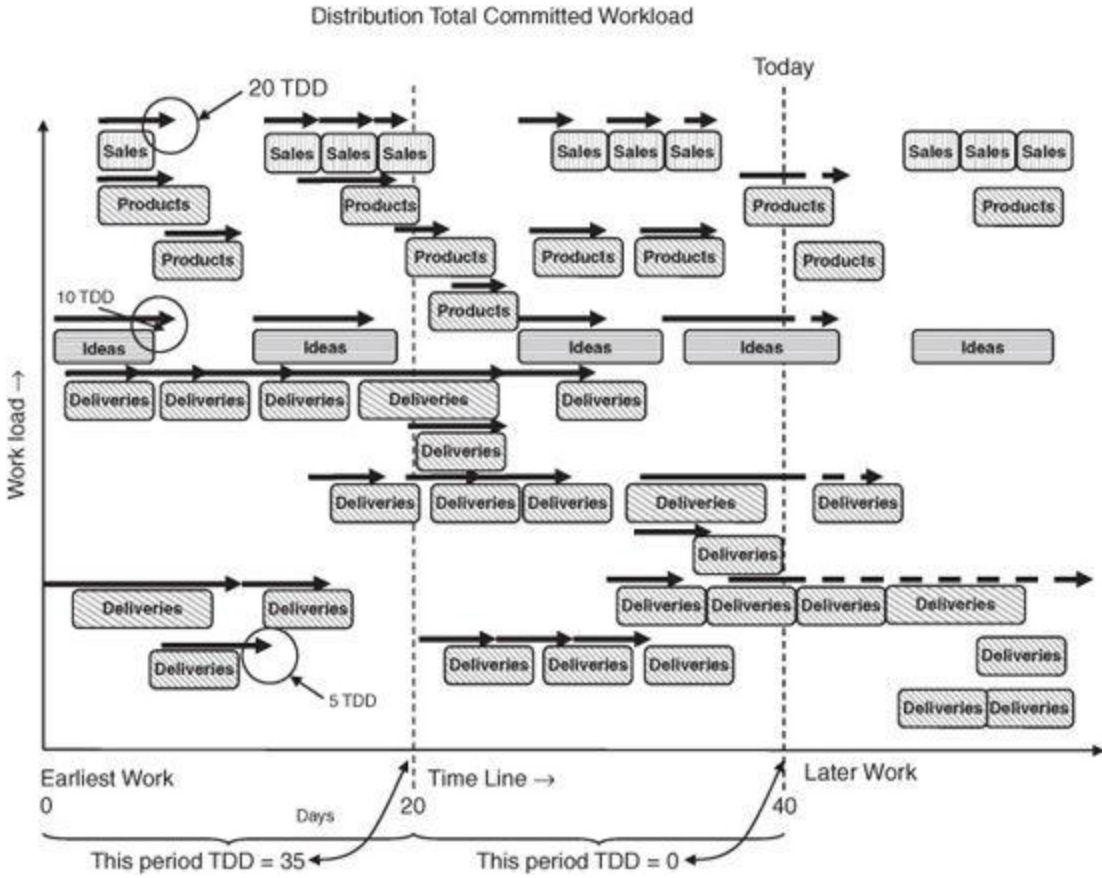


FIGURE 33-15 Distribution's total workload.

There is a third category for units or departments that have a much less direct effect on Throughput. For example, Accounting is responsible for generating monthly, quarterly, and annual financial statements. This is a critical function, so a project buffer would be maintained but the impact of this function has only a very indirect impact on Throughput. It is difficult to see how TDD would be measured for this function of the Accounting department.

These three categories of units might be classified with respect to their impact on Throughput as primary (Production, Sales, Distribution, Engineering, and similar units), secondary (Human Resources, Purchasing, IT, and similar units), and tertiary (the financial reporting function of the Accounting department). For units with primary and secondary impact on Throughput, TDD makes sense. For units or departments in the third category, it is difficult to see how TDD might be measured. However, it is well established that measures motivate performance and we would like to be able to measure units whose work falls in this third category for that reason alone. If we can develop a measure that allows senior management to monitor the performance of such units in a way that allows comparison across different units, it would be helpful.

Alternatives for When TDD Does Not Seem to Fit

Some organizational units (particularly support groups) have little control over when work is assigned and some have undetermined delivery dates. Others control their own demand and do not have delivery commitments to either external customers or other organizational units. An example of the latter is a process improvement or cost reduction group. TDD cannot be measured because there are no delivery commitments to other organizational units or customers. For such groups, the focus can be on doing

good things—generating Throughput in terms of completing whatever support tasks the unit is responsible for in a timely manner. Two examples might be helpful here. The first is the *Product Cost Improvement Program* (PCIP) at Boeing. The PCIP is a group of engineers who evaluate and implement cost reduction suggestions from various parts of the company. The group evaluates cost-saving proposals and decides which to implement. Without real delivery commitments, it is easy for the group to release many projects into the system and, due to multitasking, take a long time to complete projects. When projects are completed, however, there is a definite cost saving realized by the company. By measuring the amount of Throughput transferred through the group each day (Throughput per day or T/D), the group can see how much work is being done and track its contribution over time. Tracking the amount of Throughput delivered per day encourages the group to reduce process flow time and take actions that improve Throughput value; both increase T/D. It would be easy for the PCIP group to report T/D calculated on a monthly basis. Because the PCIP group does not have delivery commitments to other units in the organization, TDD cannot be measured for the group. However, it is still helpful to have a measure of how much the group is contributing to the organization and how well it is doing in fulfilling the purpose for which the group was created.

A second example is the hiring process of the Odessa, Texas Police Department described by Taylor et al., (2003). The department's hiring process required an average of 117 days and was not hiring sufficient officers each year to maintain the department at the desired strength. The TOC TP were used to identify what to change, what to change to, and how to cause the change, but the impetus for starting the process was awareness of the shortfall in the hiring department's production relative to its goal of hiring an adequate number of officers each year. Although the

department did not use such a measure, it is easy to see how the department could have used a measure similar to T/D to monitor its progress. In this case, “officers hired per month” would have been useful to make senior management aware of the progress of the hiring department in meeting its goal of hiring sufficient officers to maintain the required strength of the police department.

The definition of *Throughput* suggested in the *TOCICO Dictionary* (Sullivan et al., 2007, 47) in these two examples is “the rate at which the system generates ‘goal units.’” (© TOCICO 2007, used by permission, all rights reserved.) Even departments that do not have a direct impact on Throughput can define goal units to monitor their progress. Typically, T/D is calculated as the number of goal units delivered over the reporting period divided by the number of days in the period. T/D encourages groups to move quickly to complete larger Throughput value work sooner. In addition, it encourages shortening the flow time of all work, especially the lower Throughput value tasks.

Any organizational unit should be able to determine its goal units and calculate a T/D. This measure helps unit management stay focused on the unit’s goal, but some units could also benefit from another measure, described next.

Inventory Dollar Days

We can borrow another supply chain measure, Inventory Dollar Days (IDD), to help the PCIP group and the Odessa Police Department monitor their progress by focusing attention on the number of projects released into the system before there are sufficient resources to process them. The traditional definition of IDD applies to physical inventory. *Inventory Dollar Days (IDD)* is defined in the *TOCICO Dictionary* (Sullivan et al., 2007) as IDD is computed

as the sum of all current inventory on hand valued at the original purchase price times the number of days since the inventory was received by the unit being measured.

(a) measure of the effectiveness of a supply chain that measures whether the supply chain did things that it shouldn't have done, the result of which is that the supply chain is holding inventory of products customers don't want. The system should strive for the minimum IDD necessary to reliably maintain zero throughput dollar days. (© TOCICO 2007, used by permission, all rights reserved.)

Effective use of DBR and eliminating efficiency as a performance measure at non-constraints obviate the need for measuring IDD for physical inventory within the organization. However, we can redefine IDD to help with "conceptual inventory." For some groups that have no physical inventory but have a flow of small projects, assignments, papers, or mental tasks, it would be helpful to just count the number of such items in progress and measure how long they take to complete. For example, an Internal Audit department might just count the number of audits in progress and the length of time they are in progress. If five audits were in progress at the beginning of the month, and none was completed, and no new audits started, the department would report 150 "audit days" (5 audits × 30 days) incurred during the month.

Sometimes it is straightforward to assign a dollar value to the inventory of these conceptual items. In the internal audit example, the department might assign an average value of \$100 to audit hours incurred and use this to determine the inventory value of audits in progress. For example, an audit that is budgeted for 120 hours would have a total value of \$12,000. Because on average only half the hours have been incurred at any point in time, a useful simplification would be to assign an inventory value of 50

percent of the total value, or \$6,000, as soon as the audit is initiated and use this figure to accumulate IDD throughout the duration of the audit. This approach has the benefit of valuing larger audits at a proportionately higher value than small audits, with the result that IDD would be a more accurate reflection of the work that is in progress than the simple “audit-days” measure described previously.

Applying this concept to the Odessa Police Department, a value of one-half of the salary of the position being hired could be used as an inventory value. For example, if the department were starting the hiring process for a new administrative assistant at a salary of \$20,000, one-half of that amount or \$10,000 would be added to “inventory” and kept there until the person was actually hired. At the end of each reporting period, the total number of IDD would be reported. In this case, if the search for the administrative assistant had been started on the 21st day of the month, \$100,000 IDD (10 days \times 50 percent of \$20,000) would be reported at the end of the month. If the search were not completed the following month, another \$300,000 IDD (30 days \times \$10,000) would accumulate. This measure would be extremely valuable when the department is hiring a class of 10 recruits. If the average pay for a new recruit were \$30,000, IDD of 4,500,000 (50 percent of \$30,000 \times 10 recruits \times 30 days) would be incurred for each month the hiring process continues. Assigning dollar values to these types of projects, assignments, papers, or mental tasks is helpful where it is possible, because managers naturally focus on dollar amounts more easily than on simple item counts.

It should be noted that the use of IDD described here is not the same as its use for physical inventory in the supply chain. When used for physical inventory, the primary purpose of IDD is to discourage supply chain partners from doing things they should not be doing, that is, building inventory before it is necessary or building inventory that

may never be necessary. Our primary concern here, however, is not with organizational units doing things that should not be done (e.g., hiring the administrative assistant) but with providing visibility to unit management of WIP and encouraging units to strive to complete tasks quickly. The longer the projects, assignments, papers, or mental tasks are held in the system, the larger the value of IDD. There is an incentive to complete assignments and tasks quickly to stop the accumulation of IDD.

To see how IDD might apply to the PCIP group at Boeing, assume that the decision is made to release project 123A, which requires repositioning a support bracket on the 747 to save \$1,000 installation labor per plane. If an engineer starts the project today and pushes all of the necessary reviews, approvals, and other interactions with other units, the project will require 10 hours of PCIP engineer time and can be completed in one month. Without diligent follow-up and focus on this project, however, significant multitasking with its attendant switching costs and lead-time effects can occur. As mentioned previously, the PCIP group has not committed to deadlines for other units so it is up to the unit to make sure projects get done quickly. In this case, the project can easily take six months or longer to complete and may take 15 hours of engineer time if the engineer does not aggressively pursue it. During the additional five months, 25 airplanes have been completed without the improvement, costing the company \$25,000 more than they might have. Assume the present value of the cost savings over the remaining life of the 747 program is \$250,000 if the project is completed in one month. As previously, assume that half the Throughput value (\$125,000) is assigned as the inventory value from the inception of the project and that the project is started on the 26th of the month and completed on the 20th of the following month. For the first month, IDD of \$625,000 would be recorded (5 days × \$125,000) and for the second month \$2,500,000 would be

recorded (20 days × \$125,000) related to the project. Tracking IDD will motivate the engineer to stay on top of this project and push it through the system to minimize IDD in the future.

The actual results for Boeing's implementation of T/D and IDD for the 747 and 777 PCIPs are impressive. The engineers associated with the PCIPs for each aircraft began focusing on the measures and asking the right questions. Analysis of the impact of implementing these measures from June 2001 to February 2002 and from October 2002 to March 2003 showed Throughput (net cost savings) for completed PCIP projects exceeding \$62 million, a 500 percent increase over similar periods before T/D and IDD were implemented. In addition, flow time decreased by 50 percent, costs were reduced 40 percent, and the quality of the product increased. After the new measures were implemented, the group manager stated that all of the group's prior measures could be eliminated because T/D and IDD were all they really needed to know to manage their process (Mortenson 2002; Chambers 2003).

Summary of Measures

[Table 33-4](#) summarizes the previous discussion of the use of TDD, T/D, and IDD for different types of units in a complex organization.

Measure	Degree of Impact on Throughput			Units That Control Their Own Demand
	Primary	Secondary	Tertiary	
TDD	X	X		
T/D		X	X	X
IDD				X
Examples	Sales Production Distribution Engineering	Human Resources Purchasing IT Engineering	Financial Reporting Building Maintenance Legal	Process Improvement Cost Reduction Internal Audit

TABLE 33-4 Summary of Measures

Focusing for Balance (and Changing the Culture of the Company)

With TDD, the fourth Supply Chain Flow Concept is also in place for the parts of the organization that most directly affect Throughput. TDD provides the focusing process to balance flow (Goldratt, 2009). While Ford used direct observation and Ohno used the gradual reduction in the number of containers and then gradual reduction of parts per container, TDD can be used to focus on balancing flow based on time.

When using TDD on the critical process flow, it is easy to see where the system is unbalanced.¹¹ Senior management can use TDD because it has a consistent meaning across units, that is, that an organizational unit has been late on a commitment to deliver to either another unit or a customer. The magnitude of TDD is easy to interpret.

In addition, since everyone is measured in the same way, this means the top levels of the company want exactly the same thing as the bottom levels of the organization: fast, reliable flow. With this in mind, the top priority of the top of the organization is to ensure that the lowest levels of the

organization are fast and reliable. These measures link local actions to global results. Those at the bottom of the organization also want senior management to succeed at being fast and reliable because that means the bottom of the organization will be achieving its goals as well.

In contrast to TDD, T/D and IDD are more valuable to unit managers than senior management for two reasons:

1. Both T/D and IDD are unit-specific—they depend on the unit's goal and the specific types of intra-unit projects and tasks that are required to accomplish the unit's goal. Therefore, T/D and IDD generally are not comparable across units the way TDD is.
2. T/D and IDD are relative rather than absolute measures. TDD for a period is meaningful by itself, whereas T/D and IDD for a period can only be evaluated in relation to prior periods.

As seen in the example of the Boeing PCIP group, measuring T/D and IDD can have a significant impact on unit results. There is significant benefit to the units and the organization as a whole of adopting these measures even if they are not as useful to senior management due to the lack of comparability across units. Senior management, however, when evaluating individual units, would still review the trend of T/D and IDD to evaluate unit performance.

The Usefulness of Dollar Day Measures in General

Goldratt (1990) introduced TDD and IDD as measures of performance related to physical goods and the measures have been used in the TOC Supply Chain Solution. The concept of dollar days can be applied to entities other than Throughput and Inventory and can provide management useful information not previously reported. In one of Goldratt's early discussions of IDD (Goldratt, 1988a), he

compared it to common inventory measures such as inventory turnover and pointed out how IDD may be more useful to management in evaluating inventory levels. The same concept applies to accounts receivable (A/R). The normal way of describing A/R is by aging, which shows the total amount of A/R that are current, 0-30 days past due, 30-60 days past due, etc. Reporting receivable dollar days would provide management similar information condensed into one number. Another application of dollar days would be to measure lateness in paying suppliers. The payable dollar days (PDD) in this case would just be the invoice amount multiplied by the number of days it was paid late. Paying on time is critical for suppliers and PDD would provide senior management with a quick measure of how well the Finance department is taking care of vendor relationships.

A Breakthrough Injection Is Critical, but It Is Rarely Sufficient

Having everyone in the organization who has a significant impact on Throughput measured by the same simple measure (that aligns all the actions of the organization with the goals of the organization) is very important and will solve many of the problems of complex organizations. Having all units adopt T/D and IDD measures tailored to unit goals is also extremely helpful. However, two additional supporting injections are needed. The first involves Conflict Resolution and the second applies to Resource Allocation.

[Figure 33-5](#) alludes to the widespread conflicts between organizational elements within any organization where different elements have different goals and needs even though they must all work together. Having everyone who has a significant impact on Throughput measured by TDD eliminates most of the conflicts. Having a common measure

means that the goal at the top of the organization is the same as the goal at the bottom of the organization. Both senior management and unit managers want the total TDD to approach zero. This creates a new organizational culture where everyone wants the same thing and it gives a measurable understanding to the concept of balance. Those at the bottom of the organization can now feel confident that those at the top of the organization are doing the right things for the whole organization. This means job security, stability, and growth. In other words, those at the bottom of the organization want those at the top of the organization to succeed in their goals (minimizing TDD of the whole organization). Those at the top of the organization want everyone at the bottom to achieve the same goal. Cooperation happens.

Still, conflicts will exist between different elements of the organization as each part tries to improve TDD. There needs to be an effective identification, communication, and implementation tool to resolve these conflicts quickly, easily, and correctly.

Tools for Resolution

Previous chapters of this handbook have addressed the TP tools, and specifically the management skills. These include the EC, the Negative Branch Reservation, and the Prerequisite Tree (or Ambitious Target Tree). These three tools are sufficient to resolve the conflict at all levels. They work because they are not negotiation tools but tools to discover and communicate truth.

The EC addresses the combined goal (objective A) of two parties and the needs (Requirements B and C) of both sides. The conflict (D and D') comes when one side needs to act in a specific way to meet its need but this specific action impinges upon the need of the opposite side. By examining

the needs of both sides and the assumptions, a suitable injection can always be found for common conflicts.

The Negative Branch Reservation exposes how even the best of intentions can lead to negative effects. Communicating the causes of these negative effects highlights where the system can be improved. Moreover, the additional injections needed to eliminate the negative effects always improve the system as a whole. When a chronic conflict surfaces, using the EC and Negative Branch Reservation, with the parties working together, creates a new level of understanding and cooperation.

The Prerequisite Tree (or Ambitious Target Tree) is a very effective tool to overcome the obstacles facing any new initiative. Groups that work together to overcome the obstacles develop significant teamwork skills and achieve ambitious targets.

Controlled Resource Allocation

Another needed injection addresses the need to allocate resources correctly. Initially TDD will highlight areas that require senior management attention; however, once unit managers learn what the capabilities of their local resources are and how to manage effectively using DBR and CCPM, they will be making fewer commitments that cannot be met. Increasingly there will be requests for commitments from customers and other organizational units that cannot be met immediately due to capacity constraints. It will be critical for the organization to allocate resources in such a way as to maintain the balance of flow in the organization, develop the resources of the organization, and use the most critical resources in the most effective way.

In *Reaching the Goal* (2008, [Chapter 4](#)), Ricketts elegantly describes the management of the resource bench. Assigning critical resources from a central pool according to the needs

of different parts of the organization makes very good use of the resources. Managing the central resource pool to accommodate returning resources, attrition, and acquisition of resources in advance of the need is handled with a resource buffer. This resource pool concept works exceptionally well when those using the resources are encouraged to return unused resources to the resource pool as soon as they are no longer used (an IDD measure encourages this). This will only work when project and department managers know they will receive an adequate number of resources when needed.

Carrying Rickett's resource bench to the next level helps project managers and department leaders make even better use of their limited resources. Too often, the best, most qualified resources are overloaded and unable to offload work to other less qualified resources. This situation delays the development of the less qualified resources and prevents the organization from fully benefiting from the expertise of the most qualified.

The solution for this problem involves separating a small group of the most qualified resources (10 to 20 percent of like resources is sufficient) from the everyday duties of the work. This local expert group acts as a local resource bench to move in and out of the day-to-day activities as the need arises. This way, the less qualified resources can do the day-to-day activities and develop capabilities. If a less qualified resource runs into a problem that cannot be resolved within the allotted time (when TDD is threatened), then the experts from the local resource bench come and help. This develops the less qualified resource right at the time the resource is ready to learn, protects the due date, and allows a few experts to use much of their free time on improving the local processes. When on-time delivery is an absolute necessity, all resources of a group may need to participate in an all-out effort (Washington State University, 2009).

Such efforts by the experts and others create extreme teamwork.

Challenge of the Future

When complex organizations have TDD in place along with the other supporting injections, the organization is positioned to deliver in ways that other competitive complex organizations cannot match. When there are few conflicts in the organization and resources are available as needed, the organization is in a position to follow its strategy as never before. This success creates its own challenges. Rapidly improving organizations soon hit a roadblock to growth when the leadership teams are stretched too thin. Reliable leadership quickly becomes the constraint. [Figure 33-16¹²](#) shows that during periods of rapid growth, the frequency of making key decisions increases at the same time as the seriousness of each decision. Management has less and less time to make more and more important decisions. There is less and less time available for analysis and evaluation.

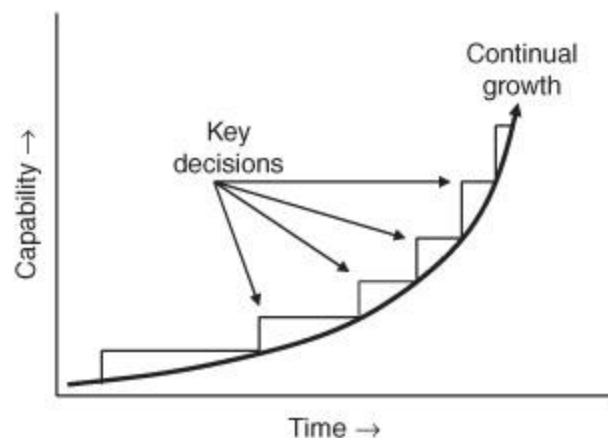


FIGURE 33-16 Decision steps under continued growth. (Used with permission of John Thompson.)

The Value of Everyone Measured by the Same Simple Measures

Under the ever-flourishing conditions of continued growth shown in [Fig. 33-16](#), the leadership teams must feel confident that they are making correct decisions and moving in the right direction. Moreover, they must feel confident that the organization as a whole can continue to accept and meet the new challenges facing them. The breakthrough injection—everyone in the organization who has a significant impact on Throughput is measured by the same simple measure (that aligns all the actions of the organization with the goals of the organization)—will go a long way toward providing the confidence needed. In a rapidly growing organization, promotions occur frequently. Those who are experienced with the TDD, IDD, and T/D measures and other TOC approaches within the new organizational culture are those best suited for leading the organization up the growth curve. However, it is often difficult to determine the effectiveness of the management team until after too many errors are made.

Leadership Certification

To solve this problem, organizations are strongly encouraged to develop or use external certification organizations to validate that members of the leadership team are aligned and all moving in the right direction at the same time. The Theory of Constraints International Certification Organization¹³ offers such certification. TOCICO maintains an online TOC Dictionary as the standard vocabulary across all functions, divisions, and companies. Many certified TOCICO members are teachers and consultants who offer services needed by complex organizations. TOCICO certification is readily available worldwide and is updated continually to current technology. Their exams meet the needs of all parts of the organization.

The most valuable element is to have most managers certified so they all speak the language (a common language), and have the same goals, measurements, and understanding of the strategic direction of the system. However, employees of any organization implementing TOC should take the TOCICO Fundamentals exam, which would ensure that they understand the basics of all TOC applications and day-to-day use of the TOC TP. Management could then be confident that all employees share a common language and understanding of management instructions and the reasons for them.

Summary

Complex organizations are composed of many individual units and departments that all depend upon each other for the orderly execution of their processes. While each unit is trying to improve and do its best, there needs to be an overall management system to get the complex organization moving and keep it moving and improving. The four Supply Chain Flow Concepts (Goldratt, 2009) set the direction for the solution. TDD, one of the TOC Supply Chain measures, provides the common measure for all units of the organization that have a direct impact on Throughput and is the mechanism for reliable and effective operation across many interconnected elements of the organization. For units and departments that do not have a direct impact on Throughput, T/D and IDD are useful for measuring progress toward unit goals. Using the three measures, complex organizations can achieve ever-increasing growth with more and more stability at the same time.

References

Barnard, A. 2003. "Insights and updates on the theory of constraints thinking processes." Paper presented at the Annual TOCICO Conference, September 9, 2003, London, England.

Chambers, P. V. 2003. *A Theory of Constraints and Six Sigma Application to Improving Cost Reduction Performance*. Pullman, WA: Engineering & Technology Management, Washington State University.

Goldratt, E. M. 1987, 1988, 1989, 1990. *Essays on the Theory of Constraints*. Great Barrington, MA: North River Press.

Goldratt, E. M. 1988a. *The Theory of Constraints Journal*, 1(3). New Haven, CT: Avraham Y. Goldratt Institute.

Goldratt, E. M. 1988b. Executive decision-making workshop. New Haven, CT: Avraham Y. Goldratt Institute.

Goldratt, E. M. 1990. *The Haystack Syndrome: Sifting Information out of the Data Ocean*. Great Barrington, MA: North River Press.

Goldratt, E. M. 1994. *It's Not Luck*. Great Barrington, MA: North River Press.

Goldratt, E. M. 1999. *Goldratt Satellite Program Session 8: Strategy & Tactics*. (Video series: 8 DVDs) Broadcast from Brummen, The Netherlands: Goldratt Satellite Program.

Goldratt, E. M. 2008. *The Choice*. Great Barrington, MA: North River Press.

Goldratt, E. M. 2009. "Standing on the shoulders of giants". *The Manufacturer*. June
http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants [Accessed Feb. 4, 2010].

Goldratt, E. M., Goldratt, A., and Ihnen, A. R. 2003-2006. TOC insights into distribution and supply chain. Goldratt's Marketing Group, <http://www.TOC-Goldratt.com>.

Johnson, J. 2009. Personal interview with Jenni Johnson, Purchasing Team Leader. August. Shippers Supply Company.

Mortensen, W. 2002. *Application of the Theory of Constraints Supply Chain Solution to the Product Cost Improvement Process*. Pullman, WA: Engineering & Technology Management, Washington State University.

Ricketts, J. A. 2008. *Reaching the Goal: How Managers Improve a Services Business Using Goldratt's Theory of Constraints*. Upper Saddle River, NJ: IBM Press.

Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*.

<http://www.tocico.org/?page=dictionary>

Taylor, L. J. III, Moersch, B. J., Franklin, G. M. 2003. "Applying the theory of constraints to a public safety hiring process," *Public Personnel Management*, 32(3).

Thompson, J. 2009. Instruction Seminar. June. Tacoma, WA. <http://www.globalfocusllc.com/>

Washington State University. 2009. The assembly game.
<http://www.vancouver.wsu.edu/fac/holt/em530/Docs/Assembly.ppt>.

About the Authors

Dr. James R. Holt is a Clinical Professor of Engineering Management at Washington State University, focusing on practical application of Organizational Behavior, Operations Research, Statistics, Engineering Economics, Simulation, Information Systems, and Constraints Management to better organizations and complex systems. He has taught Theory of Constraints principles for 20 years at WSU as a consultant, a faculty member of Goldratt Schools, and at the Air Force Institute of Technology. He is currently President of TOCICO, the Theory of Constraints Professional Certification Organization. He is happily married to Suzanne for 37+ years; they have five children and nine grandchildren.

Dr. Lynn H. Boyd has been a member of the Department of Management at the University of Louisville since 1997 and is currently Associate Professor of Management. Prior to entering academia, Dr. Boyd was a CPA with Deloitte & Touche for 14 years and also worked for the U.S. Department of Veterans Affairs for two years. He is certified as a Jonah by the Avraham Goldratt Institute. Dr. Boyd teaches operations management in both the undergraduate and graduate programs, and teaches classes in managerial decision-making and statistics.

Dr. Boyd has published articles in the *Journal of Cost Management*, *Production and Inventory Management Journal*, *International Journal of Production Research*, *International Journal of Operations and Production Management*, *Journal of Education for Business*, and

Industrial Management. Dr. Boyd lives in Crestwood, Kentucky, with his wife Rose and children Lisa and Derek.

CHAPTER 34

Applications of Strategy and Tactics Trees in Organizations

Lisa A. Ferguson, PhD

Introduction

After being exposed to writing a strategic plan using the application of the *Strategy and Tactics* (S&T) tree (a Thinking Processes tool of TOC), a Fortune 500 executive referred to his company's past planning efforts as "amateurish." How can organizations be much more effective in strategic planning and execution of that plan? This chapter explains why the S&T tree is the tool for achieving this.

In order to address how to improve strategic planning, let's first discuss the purpose of a strategic plan. This plan provides an explanation of the specific actions to be implemented over the next several years to achieve the high-level strategy or goal of the organization. Strategic plans are divided into the necessary strategies and tactics that have been agreed upon by top management.

We tend to think that strategy is what the top level of the organization focuses on, while the tactics are what the lower level of the organization implements. How do we move from strategy to tactic in our planning process? The literature does not provide clear answers on this subject. To find the answer, it is helpful to understand how this type of obstacle has been overcome successfully in the past. When Einstein

came up with his Theory of Relativity about time and space, he first had to begin by defining time. Once he realized that there was no agreement in the literature on a definition of time, he came up with his own: Time is what is measured by a clock. With this definition, he was able to develop his theory. Dr. Eli Goldratt (founder of TOC) followed Einstein's example in order to develop a theory and application for strategic planning. He defined strategy as the answer to the question "What for?" and tactic as the answer to the question "How?" With these definitions, we realize that for every action, both of these questions can and should be answered.

Strategy (S): Answer to the question "What for?"

Tactic (T): Answer to the question "How?"

Copyright © 2010 by Lisa A. Ferguson.

The S&T tree is the name of the theory and application of strategic planning in TOC. The purpose of this chapter is to provide an understanding of various applications of the Viable Vision S&T trees for organizations. A *Viable Vision* (VV) is a plan for how to become an "ever-flourishing" organization. An ever-flourishing organization is one that continues to grow exponentially, while maintaining stability at the same time.

S&T trees are a powerful tool for synchronizing all the actions needed to achieve the high-level strategy of an organization and for communication of this detailed plan to all within the organization. S&T trees can be used by any organization to achieve its strategy, not just ones focused on achieving a VV.

On Becoming an Ever-Flourishing Organization

The top strategy of the VV S&T trees is: The Company is solidly on a *Process of Ongoing Improvement* (POOGI). For a company to prosper it must be on POOGI; otherwise, competitors will just wipe out the company (eventually). What is the meaning of POOGI? Performance of the company must improve over time. Under this definition, two conceptually different curves exist—the red and green as shown in [Fig. 34-1](#). Note that each curve represents a concept and that there are multiple possibilities for each curve.

Both show performance improving. Which curve looks more realistic to you and to the people in your company? Most will answer that the green curve looks more realistic and believe that a red curve may only be possible for a short period of time for an organization. What is the real difference between the green and red curves? In a green curve, the increment in improvement each year is less than the increment the year before. The increment in absolute terms continues to increase on a red curve. Have you ever seen a company grow 5 percent (or more) year after year? This level of growth is not uncommon. Which curve demonstrates 5 percent growth per year? A red curve; the absolute growth is higher each year than the year before—5 percent of \$2 million is smaller than 5 percent of \$10 million. If you plot the performance of the U.S. economy over time, you will see that it is also a red curve. Companies traded on Wall Street must grow faster than the economy—this means that they must grow faster than a red curve, regardless of the size of the company.

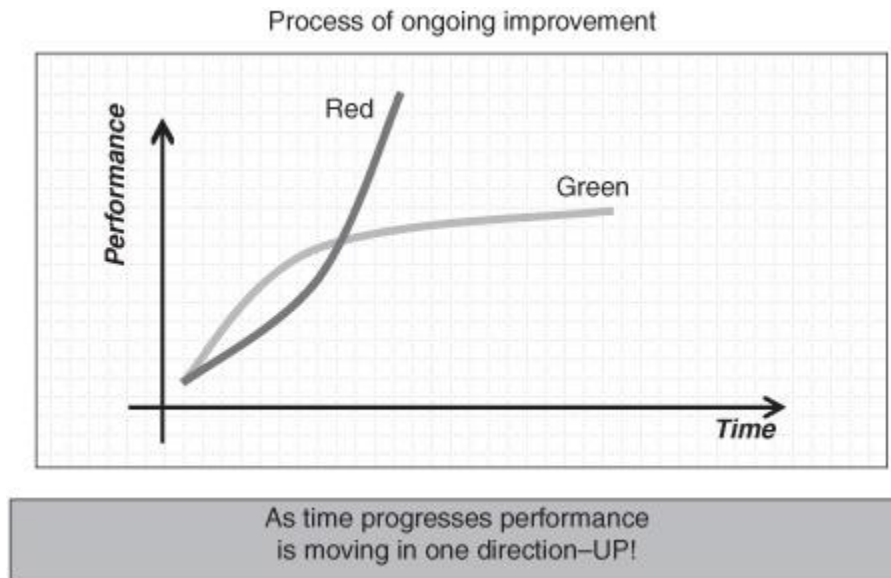


FIGURE 34-1 Process of ongoing improvement. (© E. M. Goldratt used by permission, all rights reserved. Source: Modified from E. M. Goldratt, 1999)

Why does the majority of top management surveyed think that a green curve is more realistic? This is an example of an inconsistency. The red curve represents growth. The green curve has something important for people as well—stability. People do not want the company to grow so quickly that they are spending more than 50 percent of their time putting out fires. We cannot achieve collaboration of our people without giving them what they want. People resist change only if they perceive that the change will not be beneficial. In order to not only survive as an organization, but also, more importantly, to flourish, we must achieve growth and stability at the same time.

The best proof of the need for growth and stability was written in the book *Built to Last* (Collins and Porras, 1994). Collins and Porras studied 18 highly visionary companies, also referred to as gold medalists. Some of the criteria for selecting the companies to research included being the premier institution in its industry, being widely admired by peers, having a long history of significantly affecting the

world, and being founded before 1950. They plotted the performance of those companies over time. Was it a red or green curve? It was perfectly a red curve. They discovered that the industry type does not determine whether a red curve can be achieved. Collins and Porras also compared the visionary companies to the bronze or silver medalists in their industry. Their level of growth was significantly lower than that of the visionary companies. The authors point out that two of the common factors (among others) of these visionary companies was culture and clock. A unique culture was evident in each case; after working in one of these companies for a few months, people would not consider leaving the company. Clock building is about creating a company that will continue to flourish regardless of who is leading it or the product life cycles. For each organization, you can hear the clock ticking no matter where you are in the organization—it is not about promotions or what will happen next quarter. The clock is different; the mere fact that there is a clock is obvious.

Can we ensure that our organizations are “built to last”? Ways for achieving the objectives of both red and green curves (growth and stability) at the same time were developed by Dr. Goldratt. Five different alternatives are in the public domain for achieving this objective; five generic cases of VV S&T trees cover more than 70 percent of industries that involve physical products in some form. This chapter explains the logic of the solution for each—a practical solution. The starting point is to achieve growth and stability at the same time—to build an ever-flourishing organization, not just to have a good next quarter or next year, but also to build an organization that will outlast the lifetime of a person. The top strategy of the VV S&T trees is: The company is solidly on a POOGI. “Solidly” means that we have achieved both the red and green curves together. The deeper meaning of POOGI is that the goal and necessary

conditions are achieved. All three are requirements for success. They are:

- Make more money now and in the future.
- Satisfy the market now and in the future.
- Satisfy employees now and in the future.

One of these three is the goal for an organization, while the other two are the necessary conditions (requirements) for achieving the goal. The S&T tree ensures that the actions needed to achieve all three are taken.

The Basic Structure of an S&T Tree

The structure of an S&T tree will now be explained to provide clarity before presenting a specific S&T tree. For each strategy (S), there must be a tactic (T). An S&T tree consists of a number of S and T pairs, each presented in a step. The top of an S&T tree consists of one step. Then the next level of the S&T tree below presents at least two steps (horizontal entries on the same level) further detailing the specific S and T pairs needed to achieve the higher level S and T pair, and so on until the lowest level of the S&T tree has been presented. At each level, more detail is provided about how to achieve the higher level. This is why the structure is referred to as an S&T tree. [Figure 34-2](#) provides a visual representation of the generic structure of an S&T tree.

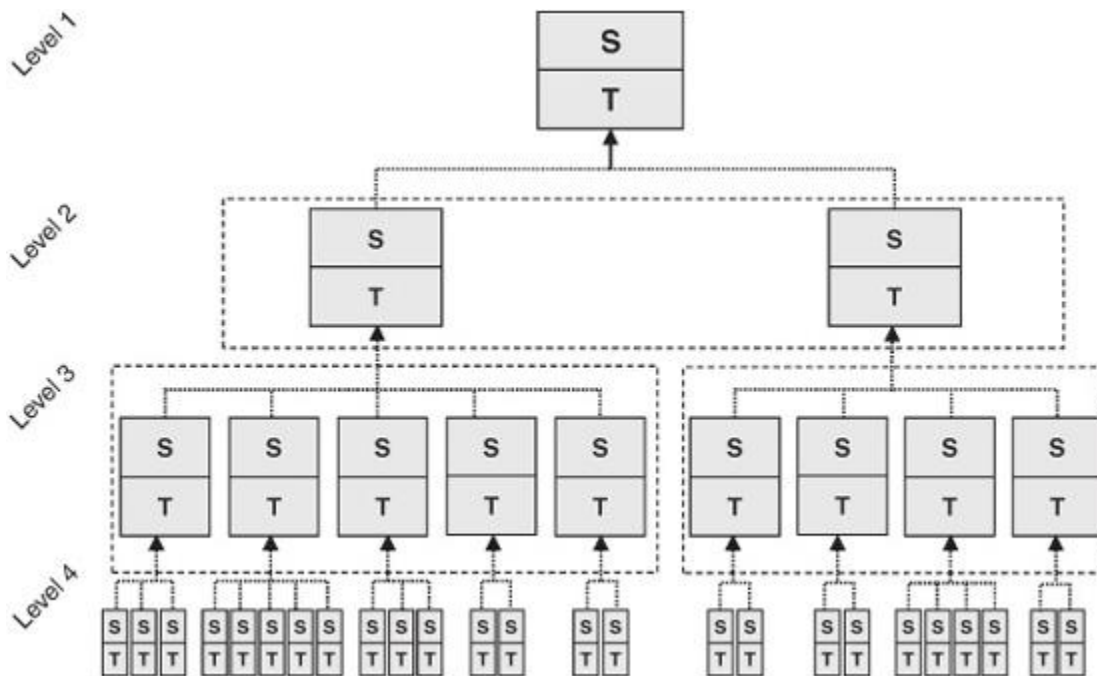


FIGURE 34-2 The generic S&T tree structure (© E. M. Goldratt used by permission, all rights reserved. Source: Modified from E. M. Goldratt, 2008).

Here is the big picture regarding the different levels of the VV S&T trees for organizations:

- Level 1 presents the pot of gold (very ambitious objective) strategy (overall).
- Level 2 presents the heart/essence of the competitive edge.
- Level 3 presents the heart of the change in mode of operation—the broad changes needed in operations and the logic regarding these changes.
- Level 4 presents the details regarding the change of mode of operation and the reasons for the change in mode as well.
- Level 5 is about how to implement the changes. It does not include the logic regarding the need to change the mode of operation; it is just about how to do the tactics we already agreed to in Level 4.

Within each step, the logic is presented connecting the parts of the S&T tree. Three types of assumptions are needed to provide the logic.¹ One is a parallel assumption (PA), the fact(s) of life, presented in logical sequence, which leads us from the strategy (S) to the unavoidable conclusion of what the tactic (T) must be. The S and T are, in effect, parallel or a match to each other. The way to read the connection is, if S and PAs, then the resulting tactic is T.

Parallel Assumption (PA): The fact(s) of life, presented in logical sequence, which lead us from the strategy (S) to the unavoidable conclusion of what the tactic (T) must be.

S || T: (Parallel Assumption Symbol).

Another type of assumption is the *necessary assumption* (NA). The NA is the fact(s) of life that explain why a specific S&T pair is needed to achieve the corresponding higher level S&T pair in the S&T tree. The NA is based on necessity-based logic, meaning that something is necessary in order to achieve something else. The NA presents the current damage of not taking the action described in the step and/or the benefits of taking the action in the step. The NA provides clear motivation for the need to take the step. The way to read the connection is, In order to achieve the higher level step, we must achieve the step below because of the NAs listed in the step below.

Necessary Assumption (NA): The fact(s) of life that explain why a specific S&T pair (step) is needed to achieve the corresponding higher level S&T pair (step) in the S&T tree.

The final type of assumption is the *sufficiency assumption* (SA). The SA is the fact of life that are common sense and commonly ignored, which if ignored will not result in all the

steps below being sufficient to achieve the corresponding step above them. An SA is based on sufficiency-based logic. With sufficiency-based logic, we need to verify that all of the components listed are sufficient or enough to achieve the result desired. However, the only way to check for sufficiency is through reality. Once all the actions have been taken, we will know if the actions were sufficient to achieve the desired objective. What we can present as an SA is guidance on what should be considered when reviewing the next level of the S&T tree as it connects to the level above. The way to read the connection is, The SA is the fact we should take into consideration when evaluating whether the group of steps below, the ones that directly connect with this step, are sufficient to achieve this step.

Sufficiency Assumption (SA): The fact(s) of life that are common sense and commonly ignored, which if ignored will not result in all the steps below being sufficient to achieve the step above them.

The Top of the VV S&T Trees

The top of the VV S&T trees, which is shown in [Table 34-1](#), is the same for the five generic S&T trees that will be discussed in this chapter. The highest-level strategy of the S&T tree is: The Company is solidly on a POOGI. The next strategy in Level 1 of the S&T tree is that the VV is realized in four years or less. The VV target for the annual *net profit* (NP) in four years is set to be extraordinarily challenging based on current thinking in business. It is believed within TOC that four years is long enough to change the culture of the company if an extraordinarily challenging target is achieved. This can be validated as more and more companies achieve their VVs. Setting a high target of NP is consistent with the research of Collins and Porras (1994),

which indicated that the visionary companies set “big hairy audacious goals.” This exponential level of growth needs to be achieved only with actions that all the stakeholders (such as shareholders and employees) of the company will agree with and support—ones that will also result in stability. This high NP target is shown to be realistic and achievable with the actions in the S&T tree and the understanding of how these actions will result in a much higher NP than previously thought achievable. For example, we can show through logic that a small increase in sales for a retailer does not increase NP at the same percentage of NP to sales ratio the retailer currently has, but rather that most, if not all, of the sales increase becomes NP because costs do not increase much, if at all.

1	Viable Vision
Strategy	The company is solidly on a P00GI. The VV is realized in four years or less.
Parallel assumptions	<ul style="list-style-type: none"> • For the company to realize the VV its T must grow (and continue to grow) much faster than OE. • Exhausting the company’s resources and/or taking too high risks severely endangers the chance of reaching the VV.
Tactic	Build a decisive competitive edge and the capabilities to capitalize on it, on big enough markets without exhausting the company’s resources and without taking real risks.
Sufficiency assumption	The way to have a decisive competitive edge is to satisfy a client’s significant need to the extent that no significant competitor can.

TABLE 34-1 Top of the VV S&T trees (© E. M. Goldratt used by permission, all rights reserved. Source: Modified from E. M. Goldratt, 2008).

The next part of this step ([Table 34-1](#)) is the parallel assumptions (PAs)—these are facts of life. PAs present the logic that demonstrates that the strategy and tactic are in essence parallel to each other. An assumption is not considered a fact of life until the people in the company agree that it is currently a fact. We read each element of the S&T tree aloud to check its validity since hearing allows us

to verify the logic using another part of our mind beyond seeing it with our eyes. The first PA is (read aloud): For the company to realize the VV, its T must grow (and continue to grow) much faster than OE. *Throughput (T) is the rate at which the company generates goal units (i.e., the rate at which the company generates money through sales, which is equivalent to sales minus the totally variable costs [TVC], such as the cost of raw materials).* In essence, this PA is stating that the company's sales must grow and continue to grow much faster than costs. The term *cost* creates confusion because the word is used with different meanings.² That is why *cost* has been defined in TOC. *Investment (I) is the money tied up in the company, while Operating Expense (OE) is all the money the company spends to generate goal units (turn Investment into Throughput).* Therefore, the cost to buy a machine is I, while the cost to run the machine is OE.

One more dollar in sales and one less dollar of costs have the same impact on NP. However, are changes in costs and sales really equivalent in the long run? The amount by which sales can increase is not intrinsically limited, while cost reduction is limited. Costs can only be reduced to zero; closing the company tomorrow will cause this to happen. Remember that we need to have the green curve (stability) as well. The two major categories of costs are employees and suppliers. Cutting costs means layoffs. Will you get the collaboration of those who remain in the company? If you lay off people after they have improved, how successful will other improvement efforts be? The other major category is the cost of purchasing raw materials for production of the physical product that is sold. It is not uncommon for a company to squeeze lower prices from its suppliers. The result is that the gross margins of the supplier are quite low—so low in fact that your supplier (which is typically small) can go out of business when the market conditions become

bad. The impact of squeezing lower prices is that the relationship between the company and its supplier is not good, but rather can be contentious. What if we instead focused on finding a way for the company and its suppliers to get both their needs met—to find a win-win solution for both companies? The result would be a better relationship and if the solution is effective, both companies will have much higher profits.³

The only way to achieve a very ambitious NP target is by significantly increasing sales. All our efforts should not be focused on reducing costs, but rather on increasing sales much faster than OE increases.

The second PA in Step 1 ([Table 34-1](#)) is not just about exhausting the cash resource. A more significant concern is about not exhausting management. Special efforts exhaust people. The result will be the inability to stay on the red curve. We also cannot afford high risks—a 20 percent risk for a decision is high if this type of risk is taken more than once. Do companies build a new plant without knowing the company will sell its capacity? Management does not know they will sell it all. Yet, they put all their cash and credit on the line. This is like playing Russian roulette with more than one bullet in the gun.

Notice how the tactic is a direct logical derivative of the PAs and the strategy. There are five components to this tactic. First, we must have (1) a *decisive competitive edge* (DCE); an edge is not based on color. Exponential growth in T cannot be achieved without a DCE. Is this enough? Technology start-ups have a DCE—they have a much better product. However, most fail within two years because they did not have the ability to (2) capitalize on their DCE. The third component is (3) competing in a big enough market. It must be large enough to sustain the growth needed to reach the VV target. Therefore, it cannot be a niche market. The last two components are about (4) not exhausting our

resources of cash and management (5) without taking real risks.

What we need to do is figure out how to achieve these five components of the tactic. The SA is also a fact of life. It is a fact, which is common sense, that most people will ignore, which if ignored will not result in all the actions being implemented that are required to achieve sufficiency. The meaning of the DCE is described in the SA of Step 1. If a significant competitor has the same DCE, you are in a price war.

After we seriously accept the definitions of strategy and tactic, we realize that we can ask these questions for every action. This means that strategies and tactics must be defined for all levels, not just at the top and the bottom of the organization. As we go down the S&T tree, more and more details for how to achieve the higher-level strategies and tactics are provided. All of the logic for the actions required to achieve the goal is provided within the S&T tree in the three types of assumptions.

There are five major splits below Level 1 of the VV S&T tree, resulting in different generic S&T trees (a generic S&T tree may need to be customized for a specific organization). Each one applies to a different environment:

- *Retailer*: sells end products directly to the client from its shelves. This type of environment is business-to-client.
- *Consumer Goods*: produces end products, but does not communicate with the client; sells through distribution networks and retailers to the client. This is a business-to-market type of environment.
- *Make-to-Order*, also known as *Reliable Rapid Response (RRR)*: produces the end item and does communicate directly with its clients; sells to another manufacturer that uses this end item as a part of their product. This is known as business-to-business.

- *Projects*: sells to client and may or may not communicate with them; what is being done is somewhat unique, though, such as a lab producing drugs or a construction company producing houses.
- *Pay per Click (PPC)*: sells final products to the client; these products, such as machines, are used by the client.

Each of these five generic S&T trees will be briefly discussed in this chapter. A full discussion of each is not possible because explaining one S&T tree fully would require many pages. Note that the full S&T trees are available in different places.⁴

The Retailer S&T Tree

We will begin by discussing the Retailer S&T tree because this S&T tree is one to which most people can relate. This S&T tree will be discussed in the most detail for that reason. This discussion will also explain generic S&T tree concepts.

All of the steps in the S&T tree below Level 1 also have an additional assumption: *the NA*. As noted earlier, this assumption explains why this step is required for achieving the corresponding step in the level above.

Level 2 of the Retailer S&T Tree

The NAs of Step 2.1 are shown in [Table 34-2](#). Each one is read aloud to verify whether it is fact in the particular retailer's environment. The first NA is read aloud: "Better availability is a consumer's significant need." Then we verify that all agree to this. Next, the second NA is read aloud: "Expecting to find an SKU and being disappointed severely erodes the consumer's impression of good availability."

Then, we can think of an example, such as a woman who finds a dress she wants but not in her size. We refer to this as a shortage. Most retailers have shortages between 5 and 30 percent of the specific products or *stock-keeping units* (SKUs) that are supposed to be available in the shop. How many sales are lost due to unavailability? Do you realize that the shortages are of the products that are the high runners—the ones that are selling well? In some cases, the customer will buy an alternative product for the one that is unavailable and may be disappointed that they had to buy a substitute. When the item is not on the shelf, sales are being lost. Retailers cannot know how many customers would have bought the SKU of which they were out of stock. Therefore, it is difficult to know how many sales are lost. After reading the third NA aloud, we point out that the forecast is not good. This results in wasting the constraint by having surpluses of SKUs. The shelf space is what limits how many different products are in the portfolio of SKUs to sell. After reading the fourth NA, we point out an additional fact not in the S&T tree, which is that a high percentage of products with a short market life are sold at markdown prices. When we present an S&T tree, we commonly provide additional information and explanation. These facts are not required to be part of the written S&T tree to reach the conclusion that the strategy that will be presented next is needed, but rather just adds to it. This NA does not apply to supermarkets. The DCE we want to achieve is to have all the products that are on the shelves be the ones that the market really wants. After reading the fifth NA, we point out that examples of this are produce, milk products, soap, and fish. Even if these items are available on the shelf, they may not be considered to be available in the mind of the customer when they are close to their expiration date. If the customers do buy ones close to the expiration date, they may decide after using or eating the product that it was not of good quality.

Notice that after reading the five NAs, it becomes clear that the resulting strategy must be the one that is stated in this step. It is important to note that there is a limit to how many times a customer will keep coming back to the same shop the more they are disappointed. The NAs in Step 2.1 (in each VV S&T tree) lead us to understand how the particular type of company addressed can achieve a DCE. The word “knowing” in the strategy is important here. It is not enough for the availability to be high; customers must be aware of the remarkable level of availability. The best kind of “advertising” in retail is word of mouth. The last part of the strategy means that parameters such as price, quality, and product selection (to name a few) must not change from their current levels. Their current levels were sufficient to be competitive until now. Therefore, not changing them, while remarkably improving availability, will result in a DCE.

2:1	Availability Competitive Edge
Necessary assumptions	<ul style="list-style-type: none"> • Better availability is a consumer's significant need. • Expecting to find an SKU and being disappointed severely erodes the consumer's impression of good availability. • Shelf space is usually the shop's constraint for better availability. A significant amount of the constraint is captured by merchandise that were ordered according to an overly optimistic forecast. • Offering many products that the market doesn't want is not contributing to the impression of availability. When the product's market-life is not long, the slow reaction time of the supply chain causes the offering to be based more on educated guesses, rather than on actual market preferences. • For a short shelf-life product, for every additional day the product spends on the shelf the customer's impression of availability deteriorates.
Strategy	A decisive competitive edge is gained by the market knowing that the company's availability is remarkably high, when all other parameters remain the same.
Parallel assumptions	<ul style="list-style-type: none"> • Besides poor quality, shortages are the main reason for a consumer's disappointment. • The current mode of operation of most supply chains, a mode of operation that is based on forecast, causes the supply chain to have a long replenishment lead time. A long replenishment time causes shortages and high inventories that block the shelf space and impair the ability to adjust the offering to the actual market preferences. • Shortages and high inventories do not only erode availability, but also (dramatically) reduce sales and increase investments. • Using TOC pull-distribution—switching to a mode of operation that is based on actual consumption—together with a proper incentives scheme to the suppliers (or better, to have suppliers that use the same consumption-based mode of operation)—ensures very high availability coupled with surprisingly high inventory turns.
Tactic	The company switches (from a forecast-driven mode of operation) to an effective consumption-driven mode of operation.
Sufficiency assumption	Building a decisive competitive edge is not easy; still, the real challenge is the ability to sustain it.

TABLE 34-2 Step 2.1 of the Retailer VV S&T tree (© E. M. Goldratt used by permission, all rights reserved. Source: E. M. Goldratt, 2008).

Now that we agree on the strategy, the question is how to achieve it. The tactic tells us how. Next, we validate whether the PAs are currently facts of life for the particular retailer. The second PA needs to be explained more. The replenishment time includes the order lead time (the time between when the first unit of an SKU is sold after an order of that SKU is received and an order for that SKU is placed

again) and the supply (production and transportation) lead times. The retailer places orders with their suppliers based on the forecasted level of demand of the SKUs. Chaos theory tells us that it is theoretically impossible to forecast accurately on the SKU level at each retail shop. We need to find a way to ensure that we do not have shortages and surpluses. The answer is quick reaction to what is selling, which is measured by inventory turns. Let's consider an example to understand the meaning and impact of inventory turns. If a retail shop currently has four inventory turns per year, then they are in essence selling what they hold on the shelves and the back storeroom in entirety four times a year. Since there are 12 months in a year, they must be holding, on average, three months worth of inventory. If we can manage the supply chain effectively to react to changes in demand, we can reduce both shortages and surpluses and significantly improve the inventory turns.

The third PA should be explained as well. The NP to sales ratio in retail ranges from 2 to 3 percent for grocery retailers to as high as 5 percent for fashion goods. The average markup on retail products from the purchase price is 100 percent. The markup is much higher for jewelry and much lower for furniture and some other types of products, such as commodities. Based on these numbers and the typical shortage statistics, we can determine how much of an impact significantly reducing the shortages would have. If shortages are 10 percent, the markup is 50 percent, and the NP to sales ratio is 2 percent, what will the NP to sales ratio become if shortages are reduced to zero? Let us assume for the moment that the costs are not affected. If sales are 100, then the TVC is 50 and T is 50. If NP is 2, then OE must be 48. If sales increase by 10 percent, then 5 more will be added to NP (half was TVC, while OE did not increase). Thus, the NP to sales ratio increases from 2 percent to over 6 percent. However, since the shortages are of high runners, it is likely that the sales will be much higher when shortages

are reduced because we cannot really know how much sales of high runners are lost when shortages occur. We will only know how much sales increase once the shortages are reduced. Even if OE does increase some, the impact on NP is still significant.

Reducing surpluses has a significant impact on investment. If up to 30 percent of the SKUs have shortages, then it is likely that more than 50 percent of the SKUs have surpluses. It is not uncommon for a retailer to have four inventory turns a year. This means that the shop is holding, on average, 3 months of inventory. If 30 percent of the SKUs are not in stock, then a high percentage must be in surplus for us to have such a high level of inventory in stock on average. The experience within TOC implementations in retailers indicates that it is not uncommon to have such a high percentage of SKUs in surplus. Reducing the surpluses affects the level of investment needed in inventory. Thus, significantly reducing the surpluses and shortages dramatically improves the inventory turns, NP, return on investment, and cash flow.

The last PA, which is shown in Step 2.1 ([Table 34-2](#)), will probably not be accepted as fact when it is read. We will ask those validating the S&T tree to accept this as fact for the time being until we can prove that it is. Assuming all of the PAs are facts, the resulting tactic must be to switch to a consumption-driven mode of operation.

The next step is to validate the SA. This type of assumption is also referred to as “Confucius says” because of the powerful common, yet uncommon, sense that is presented. What we cannot ignore when we are evaluating the next level of the S&T tree is that we not only need to focus on building a DCE, but also on how to sustain it. Therefore, Level 3 of the S&T tree needs to include one or more steps for building the DCE and one or more steps for sustaining it.

Normally, we would proceed next to validating Level 3 under Step 2.1 of the left side of the S&T tree (the left side includes Step 2.1 and all the steps under it). Instead, we are now going to review Step 2.2, as shown in [Table 34-3](#), so that we can better understand Level 2 of a VV S&T tree.

After validating the NAs of this step, we can agree that the resulting strategy must address how to expand rapidly without taking real risks or exhausting resources. The PAs in this step become facts of life after the left side of the S&T tree has been successfully implemented—meaning that all levels of the S&T tree on the left side have been achieved. In the third PA, *TPS stands for Throughput per shelf space*. The fifth PA was proven to be correct by Starbucks. They did not invest money in advertising to create a brand name. After accepting the PAs as facts, the resulting tactic must be about planning and executing a prudent expansion plan. “Prudent” is an important word here—the expansion must be done effectively without taking real risks.

2:2	Expansion
Necessary assumptions	<ul style="list-style-type: none"> • Excellent additional personnel are not easy to get. • Major expansion requires large investments and credit is not unlimited. • An established brand name in one market is not easily carried outside the boundaries of that market (into new regions or new product sectors). And it takes considerable time, money, and efforts to establish a significant brand name.
Strategy	The company rapidly expands without taking real risks and without exhausting its resources.
Parallel assumptions	<ul style="list-style-type: none"> • When operations follow excellent processes (simple, effective, and robust) it is relatively easy to train good personnel to become excellent personnel. • When a retailer consistently operates with very high inventory turns, the investment to open a new shop is considerably lower. • When a company has a recognized and established competitive edge and, as a result, all its performance (financial performance such as profit, percent of profit on sales, and ROI, as well as operational performance such as inventory turns and TPS) are much above the industry norm, the company doesn't have real difficulties in raising investments. • When a company has outstanding performance and excellent procedures, the company can successfully attract and operate a franchisee network. • Opening, within a relatively short time, a large number of shops in a given region is an effective way to create a brand name.
Tactic	The company plans and executes a PRUDENT expansion plan.
Sufficiency assumption	Considering nonexistent obstacles is almost as bad as not considering real obstacles.

TABLE 34-3 Step 2.2 of the Retailer VV S&T tree (© E. M. Goldratt used by permission, all rights reserved. Source: E. M. Goldratt, 2008).

The SA in Step 2.2 points out that only real obstacles should be considered when developing this expansion plan. Notice that many of the real obstacles to expansion have already been overcome once the left side of the S&T tree has been successfully implemented. The assumptions in Step 2.2 specifically address how these obstacles have been overcome. The right side of the S&T tree (which includes

Step 2.2 and all the steps below it) needs to address any other real obstacles that still need to be overcome.

Overview of Level 2 of VV S&T Trees

Level 2 of a VV S&T tree explains how to achieve the DCE. Step 2.1 is focused on achieving the base growth needed to reach the VV target, while Step 2.2 is focused on achieving enhanced growth. Base growth alone should achieve the NP target of the VV. To build an ever-flourishing company, more than this level of growth is required. The base growth is like getting a cake, while enhanced growth is putting the cherry on the cake; however, this cherry is much bigger than the cake itself. When the people in the organization realize that the actions implemented resulted in continued exponential growth and that these actions did not change over time (thus resulting in stability), the culture will have changed. All of the actions that are included in the S&T tree are ones that will remain in place over the long-term. For example, a change made regarding how inventory is replenished will continue in place, although more actions can be added over time to modify how replenishment is done.

Level 3 of the Retailer S&T Tree

Now we will briefly describe Level 3 of the left side of the S&T tree. Recall that Step 2.1 was focused on changing to a consumption-driven mode of operation and that the steps in Level 3 need to explain how to build and sustain this DCE. Step 3.1.1, as shown in [Table 34-4](#), explains how to build the DCE.

After reading part 1 of the PA, we point out that it is a huge mistake to push inventory into the shops. The result is shortages of some SKUs and surpluses of other SKUs. After reading part 2 of the PA, we point out that holding the

inventory back in the supply chain is more effective because the forecast is much more accurate at the distribution center (DC). Keeping more of the inventory at the DC would result in fewer cross-shipments such as between regional DCs (RDCs). Part 3 is about placing daily orders with the suppliers. Most retail executives would point out that the suppliers would not agree to this. In reality, most suppliers are struggling with handling huge orders and spikes in demand. Daily orders and frequent replenishment is a win-win solution for both the retailer and its suppliers.⁵ After reading part 5 of the PA, we point out that inventory targets should be adjusted frequently because market conditions change often.

When more than one tactic is listed, they are listed in the order in which they are implemented. These tactics need to be implemented in order to achieve Step 2.1, but they are not sufficient. How to achieve Step 3.1.1 is explained in the three steps (4.11.1, 4.11.2, and 4.11.3) that are below it in the S&T tree. The SA in this step points out a reality regarding getting stakeholder buy-in and support of a new initiative. We have to ensure that the first step in the next level of the S&T tree results in a significant and quick impact on the performance of the company in order to get this buy-in and support. The three steps below 3.1.1 are focused on implementing internal pull distribution (Step 4.11.1), the TOC solution for replenishment, keeping correct inventory levels (Step 4.11.2), and dealing with suppliers (Step 4.11.3).

Step 3.1.1 is focused on *building* the DCE by ensuring existing SKU availability. Step 3.1.2 is focused on *sustaining* this edge by further protecting and improving inventory turns. The last step under 2.1, Step 3.1.3, is focused on continuing to *build* the DCE by improving TPS by changing the product portfolio.

3:1:1	Ensure Existing SKUs Availability
Necessary assumption	The situation of almost all retailers is that in spite of constant efforts, for many SKUs the inventories are apparently too high, while for some SKUs there is no inventory.
Strategy	The company has high inventory turns; still—for each SKU—it always has enough inventory on its shelves to satisfy immediately any reasonable demand.
Parallel assumptions	When a retailer: <ol style="list-style-type: none"> 1. Holds in the shops enough inventory only for proper visual display plus what is needed for the demand (optimistically) expected within the replenishment time (transportation time from the regional distribution center (RDC)), and 2. Holds in its warehouse(s) (RDCs and CDC) enough inventory for just the expected* demand within the replenishment time, and 3. Guides its suppliers (manufacturers) according to actual daily consumption rather than batched orders, and 4. Gives its suppliers proper (monitored) incentives to improve performance (lead time and due dates), and 5. Monitors and adjusts its inventory targets according to TOC Buffer Management, 6. The retailer is able to provide very high availability while holding much lower inventories, thus resulting in high inventory turns.
Tactics	<ul style="list-style-type: none"> • The company switches its internal logistics from push to pull according to actual daily consumption. • The company implements TOC BM to monitor and adjust the target inventories in its shops and its warehouses. • The company provides daily consumption orders to its suppliers and incentives to deliver, with shorter lead time, on time.
Sufficiency assumption	To ensure an outstanding start of a major initiative, it is vital that the first substantial actions will result in immediate substantial benefits.

*Some level of paranoia (not hysteria) is recommended.

TABLE 34-4 Step 3.1.1 of the Retailer VV S&T tree (© E. M. Goldratt used by permission, all rights reserved. Source: E. M. Goldratt, 2008).

General Overview of the VV S&T Tree Structure

Let us now discuss the S&T tree generically to Level 3. The Level 1 strategy was focused on achieving both growth and stability. Level 2 has steps for achieving both base and enhanced growth. In the Retailer S&T tree, the SA of Step 2.1 focused on building the DCE and sustaining it. In other

VV S&T trees, the SA of Step 2.1 is slightly different from the one in the Retailer S&T tree in that it also addresses the need to capitalize on the DCE. This component was not needed in the Retailer S&T tree because it is enough for the customers to know that availability is now remarkable. In the other S&T trees, actions are typically needed to market and/or sell more effectively in order to capitalize on the DCE that has been built. Thus, Level 3 of each VV S&T tree consists of steps that are focused on building, capitalizing on, or sustaining the DCE.

Levels 4 and 5 of the Retailer S&T Tree

Next, we will look at the first step of Levels 4 and 5 of the Retailer S&T tree in order to understand an S&T tree better. Step 4.11.1, as shown in [Table 34-5](#), explains how to improve the inventory turns through the implementation of the pull distribution solution of TOC.

Notice that the number of PAs tends to increase, the lower we are in the S&T tree. It is important to note that the PAs are not written as a bullet list, but rather as a presentation of cause-and-effect logic. The SA in this step also reinforces the SA of Step 3.1.1.

4:11:1	Internal pull distribution
Necessary assumption	Having too little inventory guarantees a bad offering to clients. Having too much inventory (almost) guarantees a bad offering to clients.
Strategy	The company holds, in its shops and warehouse(s), relatively small amounts of inventories, which are appropriate to ensure availability.
Parallel assumptions	<ul style="list-style-type: none"> • The right inventory target is equal to consumption within the replenishment time, factored for variability. In addition, shops need to hold the appropriate amount of inventories for proper visual display. • The shorter the replenishment time, the smaller the variability is. The bigger the aggregation, the smaller the variability is (the variability in a warehouse that feeds four locations is half the variability of each location). • The replenishment time is equal to the order lead time plus supply lead time. • The conventional practices used by most retailers cause the order lead time to be significant, thus unnecessarily inflating the inventories and limiting the ability of the retailer to immediately react to actual consumption. • The conventional practices used by most retailers push the inventories into the shops (where the variability is the highest), thus inflating the inventories and limiting the ability of the retailer to react appropriately to actual consumption. • Providing the daily consumption data to the previous link reduces the order lead time to just one day and helps to prevent over-pushing inventories from the CDC to RDCs and from RDCs into the shops. • Most suppliers do not restrict the frequency of orders placed by a retailer (thus, the order lead time can be significantly reduced).
Tactics	<ul style="list-style-type: none"> • Initial inventory targets in the shops are set according to proper visual display plus optimistic expected demand during the transportation lead time from the warehouse. • The company replenishes the shops from its RDCs, based on actual daily consumption (pure pull). • Initial inventory targets in the warehouse(s) are set according to replenishment time—order, (production) and transportation lead-times. • The company replenishes the RDCs from its CDC based on actual daily consumption (pure pull). • The company orders (more) frequently from its suppliers based on actual consumption (rather than forecast).
Sufficiency assumption	An initiative should not just deliver results, but also be perceived as the cause of the results achieved; the sooner the better.

TABLE 34-5 Step 4.11.1 of the Retailer VV S&T tree (© E. M. Goldratt used by permission, all rights reserved. Source: E. M. Goldratt, 2008).

Two more steps in Level 4 are needed to achieve Step 3.1.1. Step 4.11.2 is focused on keeping the correct inventory levels through the implementation of the *Buffer Management* (BM) solution of TOC (which provides an effective priority management system), and expediting and adjusting for peak demand (as explained in the steps in

Level 5). Step 4.11.3 is focused on how to deal with the suppliers in order to achieve much more improvement in the inventory turns and the bottom line (NP, ROI, and cash flow).

Step 5.11.1, as shown in [Table 34-6](#), is the first step that needs to be implemented in retail. Because the level of sales changes over time in retail, we must ensure that we can prove that the increase in sales is the result of our initiative.

Level 5 is the lowest level of the Retailer S&T tree. Therefore, it provides the details that are required to be implemented in order to achieve Level 1 of the S&T tree.

5:11:1	Establishing reference
Necessary assumptions	<ul style="list-style-type: none"> • It is not enough that the first substantial actions of the VV initiative will result in immediate substantial benefits. These benefits also have to be acknowledged as the outcome of the initiative. • An initiative that increases sales much more than it increases expenses results in a substantial increase to the bottom line. • The variability in sales is usually high. Therefore, an increase in sales (over a relatively short period—a few months) is not indisputable proof that the VV initiative is yielding substantial benefits.
Strategy	The company realizes that the consumption-based mode of operation is a major cause of increasing profits.
Parallel assumptions	<ul style="list-style-type: none"> • When high variability exists, the way to prove the impact of an initiative is to have a control group in which the initiative is not implemented. • The control group must be representative of all shops for the proof to be valid, but should also be as small as possible because the control group will not be improved for a while. • Changes not just in sales but also in OE are compared to the corresponding changes in the control group to avoid the mistake of estimating the increase in NP based on the company's current percentage of NP on sales (the actual impact on NP is determined by the change in T minus the change in OE).
Tactics	<ul style="list-style-type: none"> • The smallest number of shops that are representative of the chain are selected to be the control group. These shops are excluded from the implementation in the beginning of the project. • The changes in T and OE for the control group and the shops, in which the VV initiative is implemented, are tracked. • Periodic, frequent reports on the results (including correct calculation of the impact on NP) are presented to (top) managers.

TABLE 34-6 Step 5.11.1 of the Retailer S&T Tree (© E. M. Goldratt used by permission, all rights reserved. Source: Modified from E. M. Goldratt, 2008).

Need for Lower Levels of an S&T Tree

It is possible that for some steps in an S&T tree, Level 6 or even Level 7 needs to be written. At this time though, Level 6 has not been written for any VV S&T trees, although we believe for some steps it would probably be quite useful to write Level 6. Another level needs to be written in an S&T tree to explain how to implement the step above *only* if it is not clear in that step how to achieve it. In the Retailer S&T tree, there are no Level 5 steps under Step 4.12.1 or Step 4.12.3, which are both under Step 3.1.2. In addition, there are no Level 5 steps under Step 3.1.3.

Details Regarding the Structure of an S&T Tree

This section summarizes what we have covered about the structure of an S&T tree with additional content now that an S&T tree has been shown and explained. The S&T tree presents the logic for how to achieve a high-level strategy. Can we agree that a strategy is an answer to the question “What for?” In other words, what is the objective? And a tactic is the answer for “How?” Brushing our teeth is an action. Can we ask what for? Yes. Can we ask how? Yes. Trying to put a strategy up high in the S&T tree and tactics all lower in the S&T tree does not make sense. In other words, it is flawed to think that strategy is for top management and tactics are for lower levels of management in the organization. Every action that we take has a strategy and tactic. Therefore, we have a number of strategy and tactic combinations or pairs, which we refer to as steps. In actuality, each tactic is an action.

The S&T tree is read from the top down. Level 1 is the top of the S&T tree, which is only one step. Level 2 is the level

below Level 1, etc. Each level of the S&T tree corresponds to a level of management in most cases. The strategy and tactic plan provides a more detailed explanation as we progress down through the layers of management. Level 1 is related to the Chairman, while Level 2 is for the Board of Directors (including the CEO). Level 3 is for the executive vice presidents (EVPs). Level 4 is for functional departments, while Level 5 is for the head of the department in the function. Level 6 is for the managers, while Level 7 (which may not ever need to be written) is for the individual employees.

We must ensure that responsibility and authority are aligned. The managers are responsible for delivering on the strategy of their assigned step. They also have the authority to change the tactics in the step for which they are responsible. For example, if the PAs are not facts of life in their company, the tactics need to be changed. We have to be careful not to have idiotic Draconian rules in the tactics. In other words, we should always check to see if the tactic we intend to implement is a logical derivative of the PAs. For example, if a tactic in the S&T tree states that 25 percent of the projects should be frozen, this rule should not be blindly applied, but rather modified in some cases to achieve the logic of the step. In some cases, it would make sense to freeze more or less of the projects in a particular company.

The S&T tree also has a sequence from left to right in a level. A step that is to the right of another step in the same level of the S&T tree cannot be implemented before the step to the left of it on the same level has *started* being implemented. The timing of when to start implementing a specific tactic needs to be based on logic. In some S&T trees, there is content in the step that points out when to start implementing the tactics.

Each step in the S&T tree includes some or all of the following statements in this order: necessary assumption(s), strategy, parallel assumption(s), tactic(s), and sufficient

assumption. The NAs explain why the step is necessary (as part of the group of steps on this level that correspond to the step in the level above) to achieve the higher-level corresponding step in the S&T tree. Therefore, there are NAs listed in each step except for the step that is Level 1 at the top of the S&T tree. The NAs need to be convincing that the action must be taken by pointing out the damage of not taking the action and/or the benefits of taking the action. The sequence of NAs is in order of the concerns that people will have. The NAs in Level 4 are focused on what is currently being done and the need to do it differently. The NAs in Level 5 are about the difficulty in doing the tactic in Level 4—the tactic we already agreed to do.

One way to better understand what an NA is results from an explanation of necessity-based logic and a visual aid. In a conflict, we understand from the *Evaporating Cloud* (EC) tool of *Thinking Processes* (TP) that a need is what is desired from a want. In other words, an action that we want to take is really focused on achieving some need. The connection between the need and the want is a necessary assumption, the explanation why this want will result in the need being achieved. It can be read as follows: “In order to achieve the <need>, we must <want> because of <necessary assumption>.” Likewise, in an S&T tree, the NA provides the connection between one S&T pair or step and another in a S&T tree as shown in [Fig. 34-3](#). It can be read as follows: In order to achieve Step 1, we must achieve Step 2.1 because of the NA of Step 2.1.

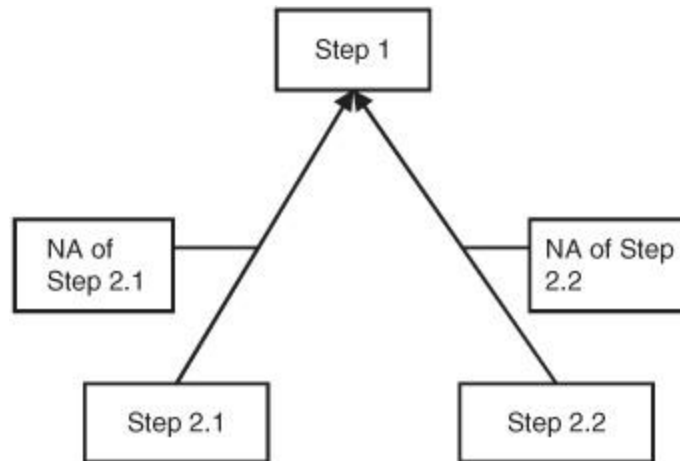


FIGURE 34-3 How the NA connects one step of the S&T tree to another.

The NAs are always generic. If the NAs we want to address are not generic, then we can devote a whole step in the S&T tree to exceptions. A step in an S&T tree can be just for a specific case. Another possible approach is to include exceptions in the PAs and the resulting tactics within a step already in the S&T tree.

A strategy is what we want to achieve. It is not stated as an action, but rather as being current reality. The parallel assumptions are for checking that the tactic will achieve the strategy. The PAs need to tell us how to do it; they are the most important part of the S&T tree because they provide the logic. The PAs explain the whole logic of the tactic—why the tactic has a real chance of making the strategy a reality. There are no surprises allowed in a tactic. The PAs must explain why the tactic is needed and will result in the strategy being achieved.

The PAs are written using cause-and-effect logic. They are written in the order that makes logical sense to reach the conclusion regarding what the tactic(s) must be. They are not written based on the sequence of the tactics. They are written so that the cause-and-effect logic is clearly presented. An entity (cause or effect) must be just one

sentence. However, we can have more than one sentence in a PA when the next sentence is a comment. It is better to keep the comment as part of one PA; it will be shorter because some content will not have to be repeated. In addition, we can put both a cause and effect in a PA. In some cases, we have freedom where to put the assumption in the S&T tree; if we have a long explanation before a conclusion, it is better to put it in PAs instead of in the NAs.

A visual aid with explanations will be useful for further explaining the logic of how the PAs are written. [Figure 34-4](#) shows a generic example with respect to sufficiency-based logic. Within the TP of TOC, we read sufficiency-based logic as follows: If A and B and C, then D is the unavoidable result. The oval shape represents a logical “and.”

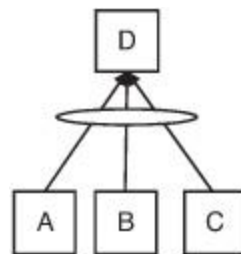


FIGURE 34-4 Sufficiency-based logic example.



FIGURE 34-5 Logic within the PAs.

[Figure 34-5](#) shows an example of how the PA is written. As stated earlier, the PA presents cause-and-effect logic, not a bullet list. If the logic were the same as presented in the previous generic example of sufficiency-based logic, then the PAs would appear as shown in [Fig. 34-5](#). Often, we would verbalize D as starting with a word such as “Therefore, . . .” to indicate that it is an effect or a conclusion that results from the previous statements of A, B, and C.

Every tactic must get results and is written as an action. The tactics are written in the order in which they are implemented. Note: The logic explaining each tactic must be clearly presented in the PAs. We can write, “Vast experience shows that . . .” in a PA if it will be explained lower in the S&T tree. The tactics in Level 5 explain exactly what to do.

It is not possible to come up with an SA that proves the steps below are sufficient to achieve the step above them. The only real test is reality. The solution was to have the SA highlight a fact that, if not dealt with by steps of the corresponding lower level group of that step, would result in sufficiency not existing. The SA must be something that is common sense, but is typically ignored; that if ignored will not result in sufficiency. They are “Confucius says” statements, which are generic. A step has an SA only if there is another level of the S&T tree written below that step.

[Figure 34-6](#) will be helpful for clarifying the logic. The SA is based on sufficiency-based logic, as described earlier. One way to read it is as follows: If Steps 2.1 and 2.2 are achieved, then Step 1 will be the unavoidable result because the SA of Step 1 was a fact of life (assuming reality verifies that sufficiency was actually achieved). Another way to read it is: If Step 2.1 and Step 2.2 are achieved and the SA of Level 1 is a fact of life, then Step 1 is the result.

When preparing to write the next level of the S&T tree, we think of the how (the titles) and the why of the steps and

the sequence of the steps. The next level below a step must have a minimum of two steps for it; otherwise, the content should be within the step itself. When creating the steps, we do not artificially break the content into two parts, but rather divide it logically into two or more parts. When I was writing Level 5 of the Retailer S&T tree, I had to figure out how to write the steps in Level 5 below Step 4.11.2 of Keeping Correct Inventory levels. I realized the way to logically divide it in Level 5 was to have one step explaining *Buffer Management (BM), the mechanism that automatically ensures correct inventory levels*, and to have another step specifically focused on adjusting for peak demand, explaining how to adjust the mechanism for upcoming expected spikes in demand that are due to a sale or other planned or known event. Later, expediting was also added as a step.

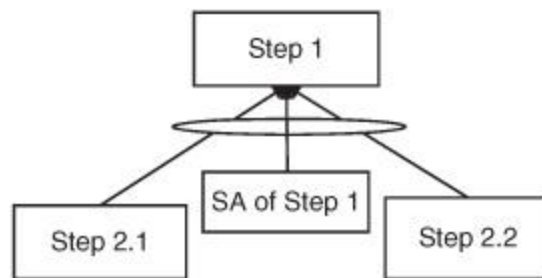


FIGURE 34-6 Visual aid for SA connection.

It is much easier to write the SA for a step after we have completed writing the steps below that step in the S&T tree; then, we have enormous intuition. The SA does not have to address all the steps below it, but it is better if it does. It is especially important to address the first step (the one to the left) of the corresponding steps below. It is better to take a known phrase/quote and change it to what we need. It makes the SA verbalization more interesting. An NA and SA can be essentially the same.

Note: We decided not to have sequence assumptions in the S&T tree. Therefore, we have to deal with any sequence that needs to be pointed out in another way. To do so, we write within a step that this step is dependent on another specific step in the S&T tree. An example of this is in Step 3.1.2 of the RRR S&T tree. This S&T tree starts with focusing on becoming a more reliable supplier by reaching at least 99 percent due date (on time) performance. It states as a note in Step 3.1.2 that the reliability selling offer should not be given the green light to start being marketed until after 99 percent due date performance has been achieved.

Key Concepts Regarding Creation of S&T Trees

We have been using S&T trees to guide TOC VV implementations for years now. VV projects are holistic TOC implementations. These VV projects consist of multiple, synchronized TOC applications (subsystem implementations), such as operations, distribution, project management, marketing, and sales. VV projects are consulting projects in which the target net profit in four years or less will be significantly higher than the best possible profit that top management of the company believes is achievable. S&T trees can be used to achieve any strategy of any system (such as any type of organization or for a person). Once an S&T tree has been developed, it is necessary to first verify that the strategy, which is one of a number of strategies we have, applies to the system (person or organization) for which the S&T tree is written. The assumptions really are facts of life. We need to validate that they are facts of life for a particular system. If they are not facts, then the corresponding strategy or tactic does not apply to the given system.

The S&T tree provides all the strategies and tactics needed to achieve the strategy in the step in Level 1. The

S&T tree provides answers to the three questions in TOC regarding how to manage effectively:

1. What to Change?
2. What to Change to?
3. How to Cause the Change?

The S&T trees that were developed to guide VV implementations provided the first significant application of TOC to answer the third question above. The development of these S&T trees clarified the steps for successfully implementing TOC. The guidance provided in Level 5 of the VV S&T trees is very useful for implementing TOC in an application/subsystem (such as a function). Level 5 provides a simplified approach to implementation that ensures that *only the injections (solution elements)* required are implemented and provides information regarding the sequence of implementing those injections.

The project network/plan for implementing the VV project consists only of the tactics in the lowest level of the S&T tree. In other words, a network showing all the dependencies (both task and resource) will be created based on all the tactics in the S&T tree at the lowest level. Specifically, if a step does not include a Level 5 while others do, then we will have tactics in the project plan from both Levels 4 and 5. This project plan in VV implementations is created using the *Critical Chain* network (which includes all dependencies and appropriate buffers) and is managed using critical chain software (from Realization Technologies).

The VV S&T trees also include strategies and tactics for ensuring that significant *negative branch reservations* (NBRs) are trimmed, thus ensuring that any significant, potential negative consequences of implementing the solution are prevented. In addition, these S&T trees include

strategies and tactics for ensuring that significant obstacles that may block or delay implementation are overcome.

The VV S&T trees are written to follow the plus buy-in process of TOC.⁶ The steps of the plus buy-in process are:

1. Agree on the very ambitious objective we desire to reach—a pot of gold.
2. Agree that reaching the pot of gold at the top of the cliff is much more difficult than we originally thought.
3. Agree that there is a direction for the solution, an anchor on the cliff against which a ladder can be leaned.
4. Agree on the solution details.
5. Overcome un verbalized fears, such as the potential NBRs of success.

The huge pot of gold is the strategy in Level 1. The strategy points out what the pot of gold is and that we can do it. It must be a strategy to which all would readily agree. Next we show how steep the cliff is (Step 2)—these are the PAs summarized in the tactic of Level 1. We show that we have nothing to hang on to in order get up the cliff. The PAs show how impossible it is to reach the pot of gold.

Step 2 of the plus buy-in process shows why it is not going to be easy to reach the pot of gold. The last PA of Level 1 usually shows why there is hope, but is not always in an S&T tree. The tactic states what will be done to reach the pot of gold, after establishing why this is the case in the PA's. The pot of gold may be a target they wanted to reach before, but decided was not possible to achieve. Step 2 is for generating credibility, to bring insight to the executives about how well we understand the problem of reaching the pot of gold. They will be more inclined to listen to us

because they will assume we might have found a solution. The PAs establish the parameters of the solution—what the solution has to address.

The NAs in Level 2 of the S&T tree are the anchor for the ladder—the third step of the plus process. The rest of Level 2 is the silhouette of the anchor. Level 2 is about how to meet the needs of the stakeholders, such as those of the external market. We consider the purpose and essence of the organization to write this level.

Level 2 below the NAs and Level 3 (which is how to build, capitalize, and sustain the DCE in the VV S&T trees) of the S&T tree are the ladder (Step 4 of the plus buy-in process)—the major rungs of the ladder and major ways in which to break your legs (Step 5 of the plus buy-in process). First, we start with showing the rungs of the ladder to climb, and then we think of major ways in which to break our legs, especially ways that can happen when we are successful.

Level 3 provides the method for achieving the strategies and tactics, but no clue how to do it. In the VV S&T trees, Level 3 explains how to build, capitalize, and/or sustain the DCE. In the VV S&T trees, Level 3 typically starts with an implementation of one of the logistical solutions of TOC (Pull Distribution, Critical Chain Project Management, or the TOC production solution of Drum-Buffer-Rope/BM). Then, if needed, a marketing or sales solution of TOC is implemented to capitalize on the DCE achieved through the logistical solution. Finally, to sustain the DCE, we are focused on how to ensure that the performance does not deteriorate as sales increase.

Every time we go down a level in the S&T tree, we are detailing the steps of the ladder. Level 4 is the level in which we are making the switch to TOC through the “golden assumptions.” These types of assumptions of Level 4 explain what we can do. In other words, we are moving from theories into practice (the actions move from being a direction to being practical actions). In reality, we tend to

call everything above the golden assumptions strategy and everything below them tactics. This is why people have the belief that the strategy is for the high level of the organization and the tactics are for the low level. This explains how the S&T tree concepts fit with conventional views of strategy and tactics.

There are cases when Level 5 does not need to be written. Here are the considerations. Are there any real difficulties to do the step in Level 4? Are there fundamental concepts that must be changed? In these cases, we need to write Level 5. We should also think about typical mistakes that might be made in implementation. We check to see if the conventional way of doing the step in Level 4 would result in mistakes. Level 5 is written if a change in a key belief is needed. Level 5 presents the actions and the implementation issues. Once Level 4 has been validated, management has already agreed to make the required changes. The logic of the how to do it is in Level 5.

The criteria for judging a solution listed below were kept in mind when writing S&T trees.⁷ These criteria are listed in the order in which they must be considered.

1. Results in excellent benefits.
2. Is win-win-win for all whose collaboration is needed. This is important because collaboration results in an increased probability of success, and a faster and more sustainable implementation. When it is not win-win-win, forces will erode it over time.
3. The risk associated with implementing the solution, multiplied by the corresponding damage, is not small relative to the benefits of implementing the solution. This is about comparing the level of risk with the level of impact. For example, if the risk is small and the potential level of damage is high, while the benefits

are huge we need to consider carefully whether we should implement the solution.

4. It is simpler than what we do now. Why is it important? The more complicated it is, the higher the chances of disillusionment of those needing to implement or support the change. If it is complicated, we do not know if it will work (if there is a chance of implementing it successfully).
5. The sequence of implementation is such that each action or cluster of actions leads to immediate, significant results, thus enabling getting everyone on board (their collaboration).
6. Does not self-destruct. If the solution self-destructs (is not sustainable), the company can be in much worse shape than it was before the solution was implemented.

All of the VV S&T trees ensure that the constraint (the factor most limiting the ability of the organization to achieve its goal) is in management's control—it is the rate at which the company can grow. The VV S&T trees are focused on achieving a DCE based on meeting a significant need of the clients. As a result, there is no limit to growth except the rate at which management will choose to grow. The S&T tree includes actions to ensure the constraint does not become internal (such as in production or sales) or a market constraint (the level of demand).

How the S&T Tree Relates to Other Thinking Process Tools of TOC

How does the S&T tree relate to the other tools of the TP of TOC? The S&T does not replace the *Current Reality Tree* (CRT)—*the map of all the cause-and-effect connecting the*

core conflict or root cause to all the undesirable effects (UDEs) in the system—and EC (conflict resolution tool) at all. The S&T tree does include some of the elements of the CRT to show why we used a different direction of a solution than the conventional approaches. The core conflict is addressed in Level 1 of the S&T tree in the PAs. The tactic of Level 1 states that we will achieve both needs of the core conflict without a compromise. The NAs of Level 2 are the assumptions that we are invalidating underlying the core conflict. The core conflicts of the subsystems are also addressed in the lower levels of the S&T tree.

The S&T does not replace the *Future Reality Tree (FRT)—the logical map connecting all the injections (solution elements) through cause-and-effect to the desirable effects (thus ensuring no UDEs of the CRT continue to occur).* The S&T tree does include all of the injections that are in the FRT.

All of the assumptions in the S&T tree must be facts of life from the CRT and FRT of the system. In other words, the assumptions must be verbalized as fact based on the current cause-and-effect logic of the system. When writing an S&T tree, it is best to conduct a full TP analysis first before writing any of the S&T tree. The core conflict, CRT, and FRT are invaluable in terms of more quickly and more effectively writing an S&T tree.

The S&T tree does replace the *Prerequisite Tree (PRT)* because the S&T tree addresses the obstacles and how to overcome them. The S&T tree provides much more logic and content than the PRT for causing the change. Two main advantages of the S&T tree to the PRT are the ability to distinguish between the big picture and various levels of detail and the ability to ensure that *chupchik (unimportant details)* are not included in the plan. This does not mean that the PRT should not ever be used. It can still be used as an effective tool for figuring out how to reach an ambitious target by determining the obstacles to reaching the target

and how to overcome them. The *Transition Tree* (TRT) is not yet replaced by the S&T tree. The S&T tree might replace it after Level 6 has been written or sequence assumptions have become part of the S&T tree. This does not mean that we would no longer use TRTs, but rather that we wouldn't need to use them when we have an S&T tree.

The Other Four Generic VV S&T Trees

Next, we will briefly discuss some key points regarding each of the other four generic VV S&T trees without presenting the steps of the S&T trees, given the limitations on how much content this chapter can address.

Consumer Goods (CG) S&T Tree

The CG S&T tree applies to manufacturers that sell to retailers.⁸ Two versions of the CG S&T tree exist: one for *make-to-order* (MTO) environments, while the other is for *make-to-stock* (MTS) environments. We will first explain the MTO S&T tree and then briefly explain how the MTS S&T tree differs. Step 2.1 of the CG S&T tree is focused on achieving an inventory turns competitive edge, while Step 2.2 is focused on achieving a TPS competitive edge. The NA of Step 2.1 is, "When most cash is tied up in inventory and availability is still an issue, improving inventory turns is a client's significant need." The resulting strategy is, "A decisive competitive edge is gained by providing a 'partnership' that delivers superior inventory turns (better availability coupled with substantially reduced inventories), when all other parameters remain the same."

The titles of the four steps in Level 3 under Step 2.1 are: Produce to Availability, Inventory Turns Selling, Expand Client Base, and Capacity Elevation. The first step is

achieved by implementing *Drum-Buffer-Rope* (DBR) and BM to improve performance in the plant. Thus, this step is focused on building the DCE. The second step of Inventory Turns Selling explains how to make an *unrefusable offer* (URO; *the marketing solution of TOC*) to prospective retailers. This step is focused on aligning the marketing and sales approaches of the supplier to capitalize on the inventory turns offer to the retailers. The third step of Expand the Client Base is about implementing the “mechanisms to generate leads, monitor, support, and effectively control their sales funnel (new clients).” Thus, these last two steps are about capitalizing on the DCE. The final step of Capacity Elevation is about ensuring that performance in the plant does not deteriorate when sales increase. Thus, this step is about sustaining the DCE.

The NA of Step 2.2 is, “When display is limited and has a major impact on sales, TPS is important to the extent that ensuring an acceptable TPS and increasing TPS are both clients’ significant needs. To rapidly achieve the VV it behooves the Company to capitalize on that fact.” The word “behoove” means that it is worthwhile to take this action although the action is not required. The resulting strategy is, “A decisive competitive edge is gained by providing a partnership that secures the clients an increase in TPS and provides a realistic chance of sharing in a much higher increase.” This means that the supplier would also benefit financially from the increase in TPS.

The version of the CG S&T tree for MTS explains how to shift from MTS to *make-to-availability* (MTA). In this S&T tree, there are three steps in Level 3 under Step 2.1: Aligning the Supply Chain, Inventory Turns Selling, and Capacity Control. The essence of the differences between this S&T tree and the one explained previously is that the changes needed in implementation differ because of how production is currently managed (MTS versus MTO).

Reliable Rapid Response S&T Tree

The RRR S&T tree is for manufacturers that sell to other manufacturers. Step 2.1 of the RRR S&T tree is focused on achieving a reliability competitive edge, while Step 2.2 is focused on achieving a rapid response competitive edge. The NA of Step 2.1 is, “When the due dates of the suppliers are notoriously bad and late delivery has major consequences for the client, reliability is a client’s significant need.” The resulting strategy is, “A decisive competitive edge is gained by the market knowing that the company’s due-date promises are remarkably reliable, when all other parameters remain the same.”

The titles of the five steps in Level 3 under Step 2.1 are: 99% Due Date Performance (DDP), Reliability Selling, Expand Client Base, Load Control, and Capacity Elevation. The first step is achieved by implementing DBR and BM to improve DDP in the plant. Thus, this step is focused on building the DCE. The second step of Reliability Selling explains how to make a URO to prospective customers (manufacturers). This step is focused on aligning the marketing and sales approaches of the supplier to capitalize on the reliability offer to their customers. The third step of Expand the Client Base is about implementing the “mechanism to generate leads, monitor, and effectively control their sales pipeline (new business opportunities).” Thus, these last two steps are about capitalizing on the DCE. The fourth step of Load Control is focused on ensuring that due dates given to clients are based on actual load in the plant. Thus, the ability to continue to meet due dates does not deteriorate as sales increase. The final step of Capacity Elevation is about ensuring that the delivery lead times are not too long as sales increase. This ensures that business is not lost due to long lead times.

Thus, the last two steps are about sustaining the DCE.

The NAs of Step 2.2 (Goldratt, 2008c) are:

- To rapidly achieve the VV, it behooves the company to have the ability to command high premiums, even on a portion of sales.
- In a non-negligible percentage of cases, the client gains heftily from rapid response.
- The client cannot get cheaper RRR (or even an acceptable alternative) from anybody except the company.
- Clients are not dumb.

The resulting strategy is, “On a considerable portion of the sales, high premiums are gained by the market knowing that the company can deliver in surprisingly short lead time.” The right side of the S&T tree explains how to implement Rapid Response. Two speeds of rapid delivery with set lead times for each are typical, with each speed of delivery having a predetermined price that is some set percentage above standard pricing.

Projects S&T Tree

The Projects S&T tree applies to companies that make a unique product.⁹ Step 2.1 of the Projects S&T tree is focused on achieving a reliability competitive edge, while Step 2.2 is focused on achieving an early delivery competitive edge. The NA of Step 2.1 is, “When the due dates of the suppliers are notoriously bad and late delivery has major consequences for the client, reliability is a client’s significant need.” The resulting strategy is, “A decisive competitive edge is gained by the market knowing that the company’s promises are remarkably reliable, when all other parameters remain the same. In the multi-projects arena, remarkably reliable (very high DDP without compromising

on the content) is defined as delivering well over 95 percent on (or before) the promised due date, while in cases of late delivery the delay is much smaller than the prevailing delays in the industry.”

The titles of the five steps in Level 3 under Step 2.1 are: Meeting Project Promises, Reliability Selling, Expand Client Base, Load Control, and Capacity Elevation. The first step of 3.1.1 is achieved by implementing the *Critical Chain Project Management* (CCPM) solution (the TOC solution for managing projects). Thus, this step is focused on building the DCE. The second step of Reliability Selling explains how to make a URO to prospective clients. This step is focused on aligning the marketing and sales approaches of the supplier to capitalize on the inventory turns offer to the retailers. The third step of Expand the Client Base is about implementing the “mechanisms to generate leads, monitor, and effectively control their sales funnel (new business opportunities).” Thus, these two steps are about capitalizing on the DCE. The fourth step of Load Control is about ensuring that the staggering mechanism of CCPM is followed even if the lead times are too long to close future deals. Following the staggering mechanism ensures that the DDP of projects continues to be over 95 percent as more project work is taken on. The final step of Capacity Elevation is focused on ensuring that the project lead times are not too long as sales increase. This ensures that business opportunities are not lost due to long lead times. Thus, the last two steps are about sustaining the DCE.

The NAs of Step 2.2 (Goldratt, 2008b) are:

- To rapidly achieve the VV it behooves the company to have the ability to win significant bonuses on many projects.
- For many projects (and more so for sub-projects) there is almost no gain in early delivery. Still, for almost every environment there are large categories of projects (less

so for sub-projects) in which early delivery brings substantial gains (sometimes the gains of early delivery dwarf the price of the project).

An example of a project that would result in substantial gains for early delivery is the opening of a retail store. The earlier it opens, the sooner revenues start coming in. The resulting strategy is, "On a considerable portion of the projects bonuses are gained."

Comparison of RRR and Project S&T Trees

Note how similar the RRR and Project S&T trees are. The NA and strategy of Step 2.1 in each are essentially the same. The only difference is in the definition of reliability. The steps under 2.1 in Level 3 are essentially the same as well. The main difference is which logistical solution is implemented as described in the first step of Level 3. Step 2.2 is similar in that it focuses on achieving more income for faster delivery. In the RRR S&T tree, higher prices are charged based on whether the delivery is rapid or super rapid. In the Projects S&T tree, bonuses are paid based on how much earlier the project is completed.

Pay per Click S&T Tree

The *Pay per Click* (PPC) S&T tree is for companies that make products that clients use. Step 2.1 of the PPC S&T tree is focused on eliminating the risk to the client, while Step 2.2 is focused on eliminating the risk to the company that makes the products. The NA of Step 2.1 is, "When a good investment is regarded as too risky, eliminating the risk is a client's significant need." The resulting strategy is, "The company gains a decisive competitive edge in large markets by providing its equipment in a way that does not involve

(almost) any risk for the client.” The titles of the four steps in Level 3 under Step 2.1 are: Market Segmentation, Market Offers Design, Pay-per-Click Selling, and Sales Funnel Management. The first two steps are focused on building the DCE, while the last two are focused on capitalizing on it.

The NA of Step 2.2 is, “Long-term profitability is not the only consideration. Additional investments and additional risks may bring a company to its knees in the short- and medium-term.” The resulting strategy is, “The additional investments needed for the PPC business are well within the capabilities of the company and the associated risks are small and manageable.” All of the steps under Step 2.2 in Level 3 are focused on sustaining the DCE. The first Level 3 step on the right side of the S&T tree is focused on implementing DBR/BM and CCPM to improve performance in the plant.

It is interesting to note that this S&T tree is the only generic VV S&T tree that does not include implementing a logistical solution of TOC on the left side of the S&T tree as the first step. Instead, it is the first step on the right side of the S&T tree.

Comparison of S&T Tree to Key Literature on Strategy¹⁰

Now that we understand more about S&T trees, we will compare this approach to the strategic planning approach that is described in the best-selling book, *Blue Ocean Strategy* (Kim and Mauborgne, 2005). The authors point out that most companies are like fish that live in red oceans. It is red from the blood of competitors eating each other. They point out that there is a way to be in a blue ocean, where competitors are not a factor. The problem is that all of their examples are based on inventions—on a customer need that was not recognized before. This is not an effective strategy

because the risk is too high. The need may not be a real need. In addition, the process for turning a need into a recognized need is not easy to do. Many companies have gone bankrupt trying to do so. We want to be in the blue ocean without the high risks. The S&T tree provides a way to achieve this. The S&T trees are focused on needs that are both real and recognized. In addition, the entire plan is focused on how to achieve the goal without taking real risks.

Porter (2008) explains how five competitive forces need to be considered when determining the strategy: established rivals, customers, suppliers, entrants, and substitute offerings. Both forces of customers and suppliers are about the power they have to pressure the company into getting what they want. Both are not relevant with respect to the S&T tree because the S&T trees provide a way to have a decisive competitive edge that no significant competitor can duplicate in the short term. The S&T trees typically entail synchronizing several functional implementations of TOC. Each implementation consists of making paradigm shifts from the traditional ways of managing. Making just one paradigm shift is not easy to do. Therefore, making more than one would be difficult for a competitor to do. Eventually, a competitor will probably be able to do so. However, the company will be prepared because another S&T tree will be ready to implement before the four years are complete. As described earlier, the S&T tree provides the win-win solution between the different links in the supply chain—between the company and its suppliers and between the company and its customers. It is important to note that the market in which we decide to have the DCE is one in which there is significant room for growth, but also one in which the company will not have more than 40 percent of the market share. This is important because the company then has room to continue to grow even if the market is going through a down cycle. The force of substitute offerings is addressed as well with this win-win

solution. Porter suggests that the way to limit the threat of substitutes is by offering better value, which is what the S&T tree does.

Porter points out that the force of established rivals can lead to price wars. The S&T trees provide a DCE that is not based on prices. In fact, in many cases the S&T trees enable charging higher prices or earning more money through bonuses based on the DCE achieved. The final force of new entrants is not really a concern either because our solution is win-win for all stakeholders. The S&T trees enable the ability to satisfy the market successfully now and in the future. Therefore, the risk of losing clients is quite low.

Porter recommends using one of three strategies: cost leadership, differentiation, or focus. Cost leadership is about being the leader in the industry based on a given level of quality. The company can choose to sell at average or below-average prices. The cost advantages are achieved through process improvements and locking in large sources of desirable materials, to name a few. The S&T tree enables the ability to achieve this type of strategy. However, it is one that others may be able to duplicate easily in a short period of time. The differentiation strategy is about developing unique attributes for the product or service that results in the company's customers valuing what they sell. This strategy is achieved by meeting significant needs of the customers. The VV S&T trees are in line with this strategy. Finally, the focus strategy is about using one of the other two strategies to capture a (narrow scope) segment of the market. This strategy is in line with the S&T trees as long as not more than 40 percent of the market share is captured.

Another contribution of Porter is the concept of value chain. Porter suggests that the company identify the key, interrelated (generic) activities of the chain and ensure each is focused on creating value. The generic core activities are inbound logistics, operations, outbound logistics, marketing and sales, and service. The S&T trees specifically address

the ability of these functional areas to enable building, capitalizing on, and sustaining the DCE. Porter argues that since a company's value chain is linked to the value chains of other companies upstream and downstream from the company in the supply chain, then the company's competitive advantage needs to depend not only on its value chain, but also on the aligned efforts of this value system. The key is to ensure a win for each link in the chain. This is consistent with the approach of the S&T trees.

The VV S&T trees ensure that the constraint, which is the rate at which the company can grow, is controlled by management. The S&T trees also ensure that the constraint does not become internal (such as within a function) or the market. In other words, the S&T tree ensures that the limit to achieving more of the goal is not the capacity of a department or the amount of demand in the market. Management has the ability to take actions to ensure that a department or the market does not become a constraint. The VV S&T trees include steps for ensuring that the constraint does not become a department or the market. These S&T trees were created with the understanding that the real constraint is management time. Having too many initiatives in the organization that management has to oversee is the opposite of *exploiting* the constraint. The usage of VV S&T trees in organizations ensures that the only initiatives are ones that will result in a significant impact on achieving the goal.

Since the S&T tree does or can address the links in the supply chain (customers or suppliers), the strategy can ensure that the constraint is not within one of these links. The winwin between the various links ensures that all the links are achieving more. However, it is possible that the constraint of the supply chain can be within one of these links. In that case, the S&T tree needs to address how to ensure that the only constraint within the supply chain becomes the ability of the entire supply chain to grow. The

focus is not just on win-win for all, but also on the understanding that unless the end customer has bought the product, no company in the supply chain has really made a sale.¹¹

Hamel and Prahalad (1994) point out that companies need to identify and focus on their area of core competence—that which proves the company's competitive strength. The criteria for the core competence are that it provides the company access to a wide variety of markets, it is difficult to imitate, and it contributes significantly to the end-product benefits. The S&T trees clearly meet the last two criteria. The usage of the S&T trees also enables meeting the first criteria as well. We have come across companies in which more than one of the generic S&T trees applies. In these cases, we can combine the S&T trees into one that is customized for them to enable achieving a DCE in more than one market and not enabling the company to have more than 40 percent market share in any one market. In any case, we usually do want to ensure that the company is not just in one market long term because the company is subjected to the ups and downs of one market. In some cases, an organization can be ever flourishing without diversification. In most cases, we would recommend that an organization plan to go into more than one market in which its core competence applies in order to reduce the risks to the organization.

Hamel and Prahalad argue that the primary killer of existing core competencies is cost cutting and silos. Neither is a concern when effectively using the S&T trees. The S&T tree does not focus on cost cutting, but rather on increasing T faster than OE increases. The S&T tree also ensures that silos are no longer an issue because the actions of the functions are coordinated or aligned to achieve the goal.

Kaplan and Norton's (1996) *Balanced Scorecard* (BSC) is a tool they developed that is utilized to translate strategy into

action. It was initially developed as a way to incorporate non-financial measures with financial measures. The BSC consists of a variety of performance measures that are divided into four categories: financial, customer, internal business processes, and innovation and learning. The process for designing the BSC for a company begins by writing the mission statement and then linking it to strategic business objectives. Next, performance measures are determined, which will be utilized to track progress on the strategic objectives. Johanson et al., (2006) point out that Kaplan and Norton think that “an effective strategic learning process requires a shared strategic framework that communicates the strategy and enables all participants to see how their individual activities contribute to overall strategy fulfillment.” This is what the S&T trees enable us to do.

The BSC has a large number of performance measures. We are aware that measures drive behaviors of people. The problem is that when there are a number of measures, it is likely that these measures are in conflict. In other words, an action taken that improves one measure hurts the performance on another. It is true that we need non-financial measures. That is why we have the three operational measures of T, I, and OE in TOC. We also know clearly what the priorities are for improving these measures when they are in conflict. In the S&T trees, there are few measures of performance. We have found from experience that when people understand what to do and how it is aligned with the goal, the right behaviors will result—assuming of course that we do not continue to use the wrong measures of performance, such as local efficiencies.¹² In addition, we argue that it is important to set up a bonus structure that rewards all employees when key performance measures of the company are improved, such as NP.

Execution of the S&T Tree

The S&T tree is a powerful tool for communication and synchronization of the efforts within the organization to achieve the goal. It is easy to learn how to read an S&T tree. The S&T tree is presented to everyone in the company to some degree. Top management must validate the S&T tree to Level 3. The validation process consists of reviewing the S&T tree to verify that each assumption is a fact of life and to deal with all reservations of management. Those who will lead the implementation of the S&T tree validate the S&T tree to Level 4. Between the presentation of Level 3 and Level 4 is knowledge transfer of the key concepts of TOC aligned with the S&T tree to be able to validate fully the logic in the S&T tree. Everyone in the company will be exposed to at least the part of the S&T tree that directly relates to them. They will also understand how their actions support achieving the goal because the S&T tree must always be presented from Level 1 down. However, it is not necessary to present all of the content of the S&T tree to do this.

The usage of the CRT and ECs in companies led to the understanding of the impact of silo thinking (each function being managed in isolation without a clear understanding of its impact on other functions or the whole system) and the many conflicts that exist within an organization. We also understood how a conflict is addressed in one silo can have negative effects on other silos. The S&T tree successfully breaks all these conflicts and ensures that all of the actions are aligned with achieving the goal.

The benefits of using the S&T tree are:

- The plan is effectively communicated to all stakeholders.

- The full logic of the strategic plan is presented and validated by the stakeholders.
- The probability of getting buy-in and collaboration of all the stakeholders increases significantly.
- Each stakeholder understands how his or her actions are directly linked to achieving the goal.
- Authority and responsibility are aligned.
- Fast results are achieved given the way in which the S&T tree is designed.

A TOC expert shared a story about a TOC implementation with the author. He stated that the implementation was done without using an S&T tree. Afterward, the S&T tree was written. He realized that a number of mistakes that were made in implementation would have been prevented had the S&T tree been written before implementation.

Summary and Discussion

This chapter provided a detailed explanation of the structure of S&T trees, with a focus on the VV S&T trees that have been released into the public domain. The discussion also covered some key concepts with respect to writing S&T trees in general. This chapter provides some guidance on how to write S&T trees. I suggest reading the article written by Goldratt, Goldratt and Abramov (2002) about S&T trees as a supplement to this chapter. Fully understanding how to write an S&T tree can be achieved through attending a workshop or reading a book written on the subject, which does not yet exist. It would have been useful to include part of the S&T tree for hospitals to show a different Level 1 and below. However, it was not possible to achieve that within this chapter. This S&T tree will be presented in some detail in materials I develop in the future

More development and usage of S&T trees has occurred within the past year. Currently, there are two types of S&T trees being used in combination in companies. At the TOCICO International Conference in Tokyo in November 2009, Dr. Eli Goldratt spent a significant portion of the first full day of his upgrade workshop discussing the S&T trees and ways to use them.¹³ The type presented in this chapter is now referred to as the Transformation S&T tree because it is effective for managing the transition of an organization from the current reality to the future reality. The second type of S&T tree, which is referred to as an Organization S&T tree, is focused on eliminating the engines of disharmony in organizations. The five engines of disharmony are:

1. Many people do not really know (cannot clearly verbalize) how what they are doing is essential to the organization. Would you be motivated if you were in that position?
2. Most people do not really understand how the work of some of their colleagues is essential to, or, at a minimum, contributes to the organization. Would you be collaborative if you were in that position?
3. People are operating under conflicts.
4. Many people are required to do tasks for which the reason no longer exists. People's intuition is always strong enough to feel it, but not always strong enough to explain it convincingly to their superiors.
5. There are gaps between responsibility and authority. You, like any other manager, know firsthand how frustrating it is to have something for which you are

responsible to accomplish, but you do not have the authority for some of the actions that must be taken.

The Organization S&T tree follows similar rules to writing the Transformation S&T tree. One exception is that each step corresponds to a person. Level 1 is the President. Level 2 includes all the people who report directly to the President, and so on. Both types of S&T trees are needed for a company to successfully become and remain ever flourishing.

References

Collins, J. C. and Porras, J. I. 1994. *Built to Last: Successful Habits of Visionary Companies*. New York: Harper Business.

Goldratt, E. M. 1999. *Goldratt Satellite Program Session 8: Strategy & Tactics*. (Video series: 8 DVDs) Broadcast from Brummen, The Netherlands: Goldratt Satellite Program.

Goldratt, E. M. 2008a. *The Choice*. Great Barrington, MA: North River Press.

Goldratt, E. M. 2008b. *The Goldratt Webcast Program on Project Management: Sessions 1-5*. (Video series: 5 sessions) Roelofarendsveen, The Netherlands: Goldratt Marketing Group.

Goldratt, E. M. 2008c. Retailer S&T tree Available at: <http://www.goldrattresearchlabs.com>

Goldratt, E. M., Goldratt R. and Abramov E. 2002a. Strategy and Tactics Tree *TOC Weekly*. December 11, 2009. www.toc-goldratt.com.

- Goldratt, E. M. Goldratt, R. and Abramov E. 2002b. Strategy and Tactics Tree: Part Two. *TOC Weekly*. December 16, 2009 www.toc-goldratt.com.
- Goldratt, E. M. Goldratt, R. and Abramov E. 2002c. Strategy and Tactics Tree: Part Three. *TOC Weekly*. December 23, 2009 www.toc-goldratt.com.
- Goldratt, E. M. and Goldratt, R. 2003a. *Insights into Distribution and Supply Chain*. Bedford, UK: Goldratt Marketing Group.
- Goldratt, E. M. and Goldratt, R. 2003b. *Insights into Operations*. Bedford, UK: Goldratt Marketing Group.
- Hamel, G. and Prahalad, C. K. 1994. *Competing for the Future*. Boston, MA: Harvard Business School Press.
- Johanson, U., Skoog, M., Backlund, A., and Almqvist, R. 2006. "Balancing dilemmas of the balanced scorecard," *Accounting, Auditing & Accountability Journal* 19(6):842-857.
- Kaplan, R. S. and Norton, D. P. 1996. *The Balanced Scorecard*. Boston: Harvard Business School Press.
- Kim, W. C. and Mauborgne, R. 2005. *Blue Ocean Strategy*. Boston, MA: Harvard Business School Publishing Corporation.
- Porter, M. E. 2008. "The five competitive forces that shape strategy," *Harvard Business Review* January 86(1):78-93.
- Thompson Jr., A. A., Strickland III, J., and Gamble, J. E. 2008. *Crafting and Executing Strategy*. 16th ed. New York: McGraw-Hill Irwin.

About the Author

Lisa A. Ferguson, PhD, is the founder and CEO of IlluminutopiaSM, an organization that is focused on “Illuminating the way to utopia for individuals, organizations and societySM.” Our Websites are located at www.illuminutopia.com and www.illuminutopia.org. Lisa is coauthor, with Dr. Antoine van Gelder, of an S&T tree for hospitals. She is currently working on writing books and papers to publish as well. Until June 2008, she spent a year working directly with Dr. Eli Goldratt (the founder of the TOC) as his technical assistant and writer (learning how to write the way he does). Since 2005, Lisa has been teaching part-time for Goldratt Schools (GS) training consultants in different countries, including India, the United States, and Japan, to be TOC Experts or Supply Chain Logistics implementers. She has a PhD in Operations Management from Arizona State University and an MBA. She taught operations management full-time in a university business school for 10 years. The last 5 years were spent teaching only MBA and doctoral students with a practical focus. Lisa has been involved with the TOC International Certification Organization (TOCICO) since its inception. She is currently a member of its Board of Directors. She is TOCICO-certified in Supply Chain Logistics Project Management, and the Thinking Processes. She resides in Sedona, Arizona and enjoys spending time with horses, hiking, and playing tennis.

CHAPTER 35

Complex Environments

Daniel P. Walsh

Introduction

At times, the challenge of making the correct decisions in a value-added chain is daunting at best and at other times it is simply overwhelming. This appears to be the case in every organization regardless of size or the complexity¹ of the products produced or services provided.

Reliance on suppliers and vendors both internal and external to our span of control further fuels the levels of uncertainty, complexity, and frustration. On any given day, we are ourselves a consumer, a producer, and a supplier of these very goods and services. Add the unreliability of our ability to forecast successfully future demand for our goods or services to the mix and it is no wonder we find ourselves mostly in a survival mode. Having observed these phenomena in many different companies within an industry sector and indeed across multiple industry sectors, the survival mode appears to be common practice. So much so that it is accepted and viewed as a fact of life that cannot be easily changed in spite of significant investments in improvement initiatives (Brown et al., 1994), and the environment has only grown more complex since this article was written.

If we view our organization as a system, then by definition all of the activities are connected. At first glance, they may

appear to be independent of each other but in reality, any action taken by one of the activities will impact the others. Then it follows any real and lasting change for the better must be based on a systems approach; all changes must not just improve a local activity, but rather the entire organization. There are two characteristics of all systems (Goldratt and Cox, 1984): dependency of variables or activities and fluctuation (more commonly referred to as variability). Even if this tenet of improvement is recognized and accepted, it will immediately create a conflict with existing metrics. This conflict highlights the requirement for an overarching common set of metrics for evaluating individual contributions of all activities while establishing connectivity to the performance of the organization as a whole. Once these new metrics are in place and effective management tools assessing the impact of internal and external variability are being used, then the variability can be evaluated quickly and corrective action taken to protect the organization's performance.

Copyright © 2010 by Daniel P. Walsh.

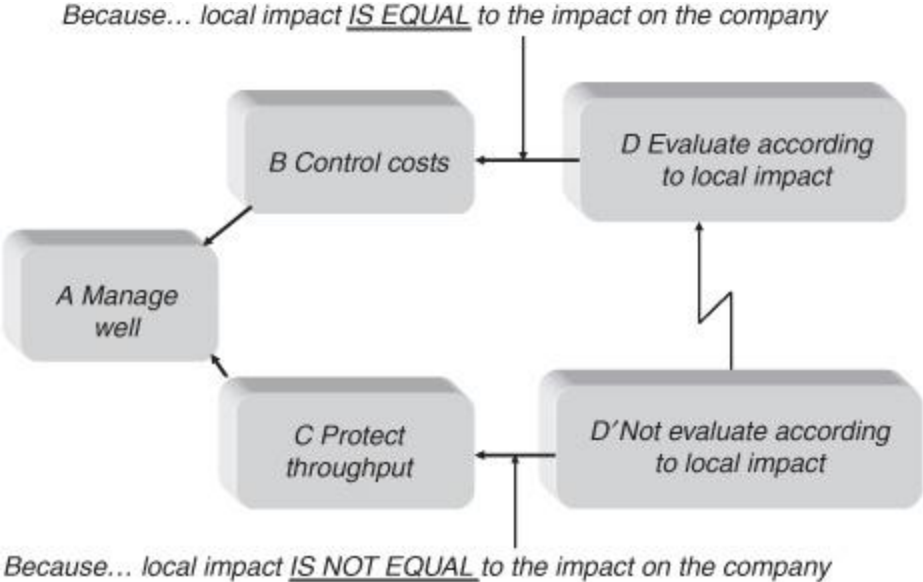


FIGURE 35-1 Evaporating Cloud of managers' dilemma of judging the system performance. (© E. M. Goldratt used by permission, all rights reserved. Source: E. M. Goldratt 1999. Viewer Notebook 137.)

The purpose of this chapter is to provide a better understanding of why addressing the effects of local variability is crucial to developing strategies that are more effective for managing entire supply chains. Again, these new metrics and approach must provide connectivity from the local activities to the global Throughput of the organization. In addition, it is important to use the correct planning, scheduling, and controlling algorithms; in other words, make sure the right tool is being used. Lastly, it is important to make sure these tools and algorithms are holistically employed.

Brief Background

First, we must better understand the chronic dilemma virtually every manager faces on a daily basis. To illustrate the dilemma and the resultant conflict, we will use a simple *Evaporating Cloud* (EC) developed by Dr. Eliyahu Goldratt (1994; see [Fig. 35-1](#)). In order to [A] manage well we must [B] control costs; in order to [B] control costs we (managers) must [D] evaluate and make decisions based on how it impacts locally. The other side of the dilemma is that in order to [A] manage well we must [C] protect the company's Throughput and fulfill our commitments to the market; in order to [C] protect the company's Throughput, we [D'] must not make decisions based on local impact. The needs of the company, [B] controlling costs, and [C] protecting Throughput, are necessary conditions and must be achieved in order to [A] manage well. The conflict is very clearly defined as being between whether we [D] make decisions

based on local impact or we [D'] do not evaluate according to local impact.²

Now why do we feel compelled to evaluate according to local impact? We feel compelled because of the ingrained assumption that the local impact of decisions is equal to the impact it will have on the company as a whole. In fact, this is consistent with common business folklore and is fortified by what is accepted and being taught in virtually every learning institution throughout the world.

The other side of the dilemma is that in order to protect Throughput, many times we must not evaluate and make decisions based on the local impact, rather do whatever it takes to meet our commitments to the market. This, of course, is the familiar phenomena commonly referred to as "firefighting," the bane of all managers. It also manifests itself by managers focusing on local metrics during the first part of a reporting period and then later in the reporting period shifting the focus to meeting orders whose due dates are starting to slip. When this occurs, the focus is no longer on local impact, but rather on delivering our products to clients.

It is clear this dilemma must be addressed or managers at all levels will remain frustrated and the true potential of a company will never be achieved.

Guiding Strategies

If this dilemma is the starting point, then we have two broad guiding strategies available:

1. The first approach is focusing on improving the individual parts of the organization within our span of control as "fires" crop up. This approach has been the predominant approach and continues to remain popular among many process improvement

practitioners and managers. It is based on inductive reasoning and the belief that improving individual parts of the organization will result in improving the performance of the organization. Starting with the early pioneering efforts (see for example, Alford, 1934, sect. 4), most of the literature and developments on organizational improvement (Churchman, 1968) have focused on this piecemeal or fragmented approach. Indeed, the majority of the widely used tools and methodologies (see, for example, Zandin and Maynard, 2001; Barnes, 1980) can trace their origins to these scientific management tenets (Taylor, 1911).

2. The second approach views the organization in its totality, focusing on a systems approach (Churchman, 1968) for improvement. It is based on deductive reasoning, long a cornerstone for breakthrough advances in the sciences and starting to show considerable promise in some of the more advanced evolving business methodologies (Rummler and Brache, 1995).

Today there are many powerful tools and methodologies available such as TOC, Lean, Six Sigma, Business Process Reengineering, etc. to help implement these improvement strategies. Still, the results have been mixed. In some cases, improvements have been documented; in other cases, the organizations showed little improvement or even none at all. Even when initial improvements were achieved, the sad reality was many of those were not sustainable. Almost invariably, the improvements took longer and were more difficult than expected.

So, where does that leave us? Rather than attempting enhancement or improving existing tools and methodologies, we can focus instead on how to holistically develop and employ a significantly more effective solution

set. This focus will require building on and leveraging the current available body of knowledge. Perhaps it would be helpful if we first gained clarity and understood why many improvements fail to meet managers' expectations.

The limitations of following the strategy of improving the individual parts of the organization lead to managing the individual parts in isolation. If everyone is managing their areas of responsibility this way, then all of the fine tools and methodologies are focusing on improving the individual parts separately. Local effects reflect the impact of problems that exist within the system of operation. Measurement of these effects on isolated "local activity performance" does not necessarily lead us to understanding the systemic problems that may be leading to negative performance. We can all agree, then, that the improvements can be summarized as follows:

$$I = \sum_i^n i_1 + i_2 + i_3 + \dots + i_n$$

where I is the sum of the individual improvements i from 1 to n improvements.

Now, it is important to accept the painful reality that the sum of the individual improvements has very little to do with improving the performance of the organization (Goldratt and Cox, 1984 [Chapter 4](#); Johnson and Kaplan, 1987; Goldratt, 1988) and will be a pure random event if it leads to any improvement of the organization. The sum of individual improvements is simply the summation of disconnected events. This can also be described as sub-optimization. It appears that this erroneous assumption—action taken locally will necessarily result in improving the performance of the organization—is one of the main contributing causes for failure to achieve real and sustainable enterprise improvements. Therefore, it must follow that this erroneous assumption must be challenged

and *de facto* abandoned, replaced by an approach focusing instead on improving the performance of the enterprise.

In order to develop an alternative approach, we must focus on improving the performance of the individual areas, such as a department or an area of activity, while providing connectivity for improving the enterprise. In other words, we must improve individual areas only if we can establish a cause-and-effect relationship between the two, showing that local improvement translates into global improvement. This will require a fundamental shift in our thinking. Before we pursue this line of reasoning, first we must agree that in any enterprise there are two indisputable and absolute truths:

- Every function and task within the enterprise is connected and therefore its outcome will affect other parts of the enterprise. Therefore, regardless of the complexity we must understand the cause-and-effect relationships the functions and tasks have on the individual parts and more importantly on the performance of the entire enterprise.
- Every part of the enterprise is subject to uncertainty, which is simply another way for describing the inevitable variability experienced in actual execution. Regardless of how meticulous our planning and scheduling, when actually executed, uncertainty and variability will inevitably affect our efforts.

These two tenets (dependent events and statistical fluctuations) provide the foundation for developing any breakthrough holistic approach (Goldratt and Cox, 1984, [Chapter 15](#) and 17), leapfrogging our ability to significantly increase the Throughput (discussed later) of a single enterprise or the larger value-added processes of an entire supply chain. Another important element of this new holistic approach is providing relevant performance and operational metrics to monitor the stability of the enterprise on a day-

to-day basis. These metrics must provide connectivity on the short term, highlighting when and where specific action must be taken while providing longer-term visibility for effective risk management.

An operational metric must have a cause-and-effect relationship providing connectivity between an action taken and the positive or negative impact it will have on the organization's Throughput. Therefore, in most cases if these operational metrics are providing the priorities for managing, they are focused on increasing Throughput. As this is done, the recurring costs shall not increase and any variable cost increase will be significantly less than the corresponding increase in sales. An example of an operational metric is using the speedometer in an automobile while driving on a trip. The output of the speedometer is the effect of the input of how hard we press down on the accelerator pedal. Therefore, if we have calculated the average speed that must be maintained in order to complete the journey on time, the information received from the speedometer will allow us to take the correct action. This is in real time, not in hindsight. In this example, a performance metric would be measuring on the road map the distance covered during the journey. Measuring the variance of the distance covered vis-à-vis what was expected to be covered is important but of very little use in making real-time operational decisions.

Throughput Accounting

The *Theory of Constraints* (TOC)³ defines *Throughput* (T) as *Sales \$* (S) minus *Truly Variable Costs \$* (TVC). It should be pointed out that all recurring costs including fixed labor costs are captured as *Operational Expenses* (OE) (Corbett, 1998).

If decisions are being made using operational metrics and they are focused on increasing Throughput, then it is possible to have the organization's financial metrics aligned as well (Corbett, 1998). TOC builds on this concept and recognizes that an organization is a system and therefore regardless of how well it is managed, its ability to increase Throughput will be limited by the system's constraint. Furthermore, if we have identified what and where the constraint is and we are subordinating everyone's efforts toward maximizing its effectiveness, then we have unlocked the secret for maximizing the organization's Throughput (Goldratt and Cox, 1984).

As we can see in [Fig. 35-2](#), we now have a model for resolving the conflict depicted in [Fig. 35-1](#), which most organizations face on a daily basis. The conflict of course is whether to take action in order to control costs or take action to protect Throughput. It is important to note that this conflict is in large part caused by using performance metrics to evaluate individual parts of the organization rather than using operational metrics to evaluate contribution to Throughput. This is analogous to driving your automobile by looking in your rear view mirror and using the history of what is behind you (performance metrics) to guide future decisions ([Fig. 35-1](#)). On the other hand, change to a new model by focusing on looking out the front windshield (operational metrics; see [Fig. 35-2](#)).

We are now using the same measurements to make operational and financial decisions. Once this new model is adapted, it is very easy to turn the operational metrics into performance metrics. Since the new model is measuring the rate of Throughput being generated at the constraint, it simply requires adding up the individual contributions at periodic intervals.

$$T = \sum_1^n i1 + i2 + i3 + \dots + in$$

where T is the periodic Throughput and t is the individual contributions to T from 1 to n .

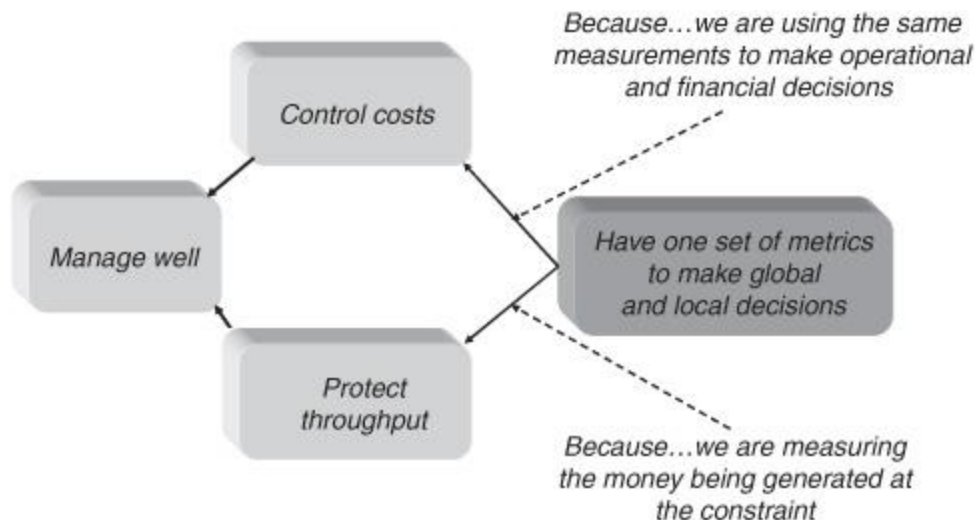


FIGURE 35-2 Evaporating Cloud solution of managers' dilemma of judging the system performance. (© E. M. Goldratt used by permission, all rights reserved. Source: Modified from E. M. Goldratt, 1999.)

The following discussion is intended to provide a roadmap for developing such an approach. It is important to share that this approach has been successfully employed across organization types and different industry sectors. I believe it has universal applicability in private and public organizations.

A Holistic View

In order to develop a holistic approach to better achieve the goals of the enterprise, it stands to reason we must first model our value-added chain as one system. Before we discuss the modeling, we must address the characteristics of a system.

There are two characteristics present that can be used in describing any system:

- Everything within the boundaries of the system is connected, which means all of the elements are subject to cause and effect. None of the elements operates in isolation. At first glance it may seem so, but one must continue looking until the connectivity is established. [Figure 35-3](#) provides a top-level system depiction of a typical company that is part of a much larger supply chain. As the systems architecture for the company is developed, a much more detailed view will be modeled. The interdependencies will be identified and the flow of information and work that culminates in a value-added product or service emerges. This is a very important part of the planning process. This is a precursor to developing the approach of dealing with variability, which is the execution part of our systems architecture.
- In execution, the individual elements are influenced by variability (see [Fig. 35-4](#)). Therefore, due to the connectivity of the elements, the variability is transferred throughout the system and thus affects the outcome of the system itself. Since variability can never be eliminated, an important part of the systems architecture design must include the capability to better manage and mitigate the variation.

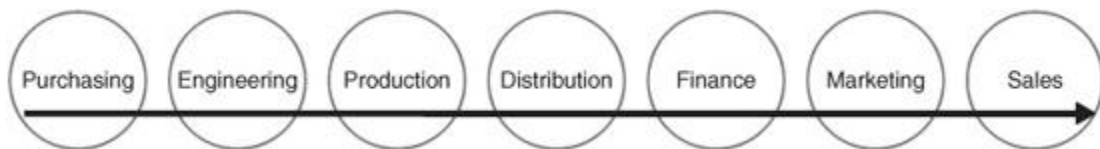


FIGURE 35-3 Links in an organization chain.

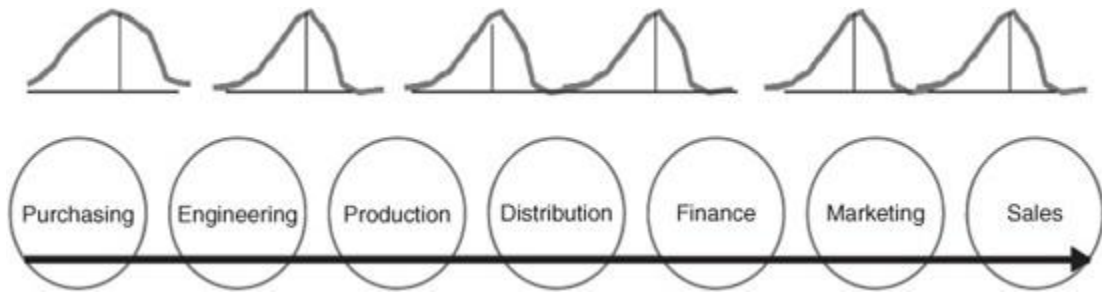


FIGURE 35-4 Statistical fluctuations and dependent resources.

The majority of companies find it very difficult to achieve their planned objectives, so much so that it leads to a belief of inevitability, of not being able to control the constant stream of uncertainties that confront them on a daily basis. Focusing on individual performance levels instead of the performance of the enterprise seems to be the only choice. The uncertainty that causes the fire fighting is actually the manifestation of variability. The negative impact to the company is a result of not being able to mitigate the impact caused by the inevitable variability but having to respond in a fire-fighting mode.

Categories of Variability

There are two categories of variability—common and special cause.⁴ They have different origins but both can adversely affect the performance of the company. Many different kinds of management philosophies have evolved attempting to minimize their impact. For example, in Lean, the kanban is used to signal variation; when it appears, it will start chocking the release of work to control the amount of *work in process* (WIP). This recognizes the fact that if work is authorized and released prior to resolving the cause of the variability, the queue increases, which will increase cycle time. Similarly, in Six Sigma, process control charts such as

X bar and R charts look at specific process variability. This in turn highlights areas for improving an individual process while providing feedback in execution.

Many companies may see improvements with these approaches to managing variability. However, today there is growing consensus that something additional is needed to get them to the next level. There is a need for an additional classification of variability that will provide greater understanding and insight in choosing the correct planning, scheduling, and execution applications that will provide better focus. First, it is important to examine the confusion and negative impact that is being caused and view it in a historical context.

Tools Selection

In every organization, there is a requirement to plan, schedule, and execute a series of actions in order to provide a product or service. Specific TOC applications and tools have been developed to manage different parts of the organization such as departments, work centers, etc. The three predominant application solutions for planning and control systems⁵ are:

- *Project Management System*—This is used to manage the projects in the company.
- *Production Planning and Control System*—The origin of this application is in manufacturing. However, it has evolved into many other parts of the company such as the customer service and administrative areas.
- *Material Management and Inventory Control System (supply chain system)*— Primarily focused on material procurement, transportation, warehousing, and inventory control.

Which application solution the organization uses is determined by the product or services provided to the market. It is interesting to note that many companies locked in old paradigms using production planning and control systems such as MRP and MRPII, measures, critical path project management and distribution systems such as DRP and DRPII have not taken advantage of the evolving thinking and technologies now available and are thus blocked from adapting them. Building on the emerging thinking and new tools available, companies are already leveraging on the conclusions reached by Schragenheim and Walsh (2004) that deeper understanding of when to use each of the logistical tools, the application solutions for Project Management, Production Planning and Control, and Material Management and Inventory Control will lead to powerful hybrid solution sets. An example will be shown later in the chapter. In fact, companies using holistic planning, scheduling, and execution techniques such as the Integrated Enterprise Scheduling engine which focuses on holistically managing the value-added chain Throughput rather than Throughput of the individual parts, are obtaining remarkable and sustainable results. The rationale and explanation for such an approach was highlighted in an article written by Schragenheim and Walsh (2004). Indeed for the first time it appears there are software solutions being developed recognizing the requirement and immense potential in being able to better manage the negative impact on the enterprise caused by the inevitable variability in execution.

A Closer Look at Variability

So let us take a more in-depth look at how variability affects an enterprise value-added chain. [Figure 35-5](#) depicts workflow in an organization. The time to complete each task

consists of set-up time plus work time plus set-down time plus resource queue time. In the case of a typical shop floor scheduling routing, this represents three different resources processing three individual tasks on a single part. In the case of a project network, this represents three different resources processing three tasks that are supporting a single part, or it could be tasks that simply receive information or results from the predecessor task in order to work on the successor task. The material required to support these tasks will be procured, placed in inventory until transported using an entirely different schedule. This is a very complex effort indeed.

The queue time or white noise is the time when no value-added or productive time is being realized. In other words, [Fig. 35-5](#) is showing how the elapsed time is still being accumulated between the productive tasks while providing no value-added to the output of the organization. The accumulative effect of productive time plus queue is equal to the total cycle time. Therefore, it follows that for any organization to improve, any scheduling algorithm must be able to synchronize and leverage the availability of the resources in order to eliminate this excessive idle time to maximize the Throughput of the organization. This means that to increase Throughput, which was previously defined as sales minus TVC, the organization must accelerate the flow of work or, more precisely, the *rate of work flow*. That is, with a given amount of resources, the organization must be able to deliver the final product to the market sooner. This rate of workflow is the key to being more responsive to the customers and increasing the company's profits.

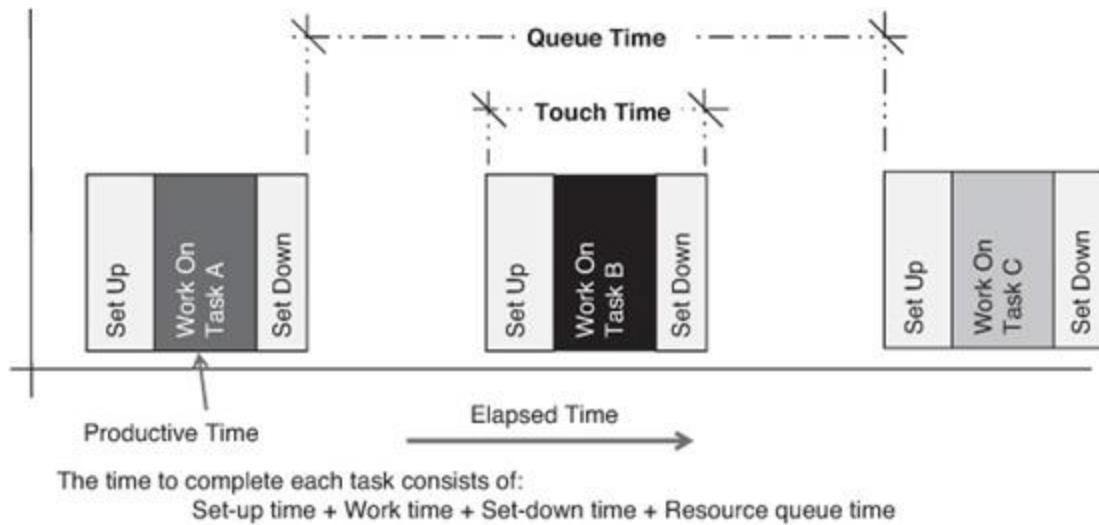


FIGURE 35-5 Elements of task time for a dependent series of tasks.

The company must find ways of reducing the cycle time of producing and delivering its products. The easiest and most effective way is by reducing the queue time ([Fig. 35-5](#)) of resources waiting to be utilized, the principle reason WIP starts to increase, which leads to resources being significantly more productive (Little's Law; Hopp and Spearman, 2000). There are different ways this can be accomplished but one fact is indisputable—every company is susceptible to variability; therefore, any successful solution must be able to better manage the uncertainties, changing of priorities, and schedule changes. This variability causes major impact to the work schedules, which become the single greatest contributor to the resource queue. There is a direct relationship between resource queue time and productivity. So if cycle time must be significantly reduced, any breakthrough scheduling algorithms *must* reduce variability when possible, thereby reducing queue times. This must be done while providing real-time information to managers to mitigate the increased risk of managing variability.

The uncertainty of resources, material availability, and required technical information, subjected to the effects of unforeseen common and special causes, can be expressed as variability. Lacking a deeper understanding of how the variability can be mitigated leads to many instances of companies actually using the *wrong* scheduling tool while attempting to better manage the variability. For example, they may be using only project management tools when this may not be the *best* scheduling algorithm because they view themselves as a “project management” company, when in reality different parts of the company may be subjected to different kinds or types of variability, which means more than one scheduling algorithms is required. Alternatively, they may only be using production planning and control scheduling tools because they view their company as a floor scheduling/manufacturing company. In addition, regardless of which tool they decide to use, in many cases they may decide to schedule and manage their material requirements by imbedding them in whatever work-scheduling tool they are using instead of using the appropriate material management algorithm.

Many of these mistakes can be traced to a lack of understanding of the origin and the cause and effect of the variability. In order to clarify this confusion, a further classification of variability is required. This will help companies decide which algorithm is appropriate and lead to more effective planning, scheduling, and management of their environment.

There are three different types of variability significantly affecting an organization.

- **Type 1**—*This occurs when most of the variability is within the task itself and not in the resource queue (see [Fig. 35-5](#)). The most significant known or anticipated variability will be in the work being performed in task A, B, and C. Remember in planning we are identifying what*

work must be done and the resources required within the tasks. This is not implying that in execution there will not be variability due to lack of resources, or set-up or set-down time. In fact, it is very highly likely that many of the tasks will be impacted by the required resources not being available. Conversely, some of the resources will spend time in the queue waiting for predecessor tasks to finish, thus allowing the successor task to start.

- **Type 2**—*This exists when the variability within the task itself is relatively low and most of the variability is in the queue.* This assumes well-defined manufacturing processes and well-defined tasks. In [Fig. 35-5](#), Tasks A, B, and C the variability is low because this particular work or something very similar has been done many times before. In companies using MRP or MRPII, the manufacturing routings are readily available and will be incorporated into the master schedule. The same can be said for the set up and set down; the required time is well known and variability is minimum.
- **Type 3**—*Occurs when the variability is in the demand pattern of material requirements.* This can be within the company if the part is currently in inventory or is a component being manufactured internally. Sometimes the material is outsourced and must be delivered in time to support the company's master schedule. This is further complicated by having to anticipate future market demand for all products, which of course determines what material is needed, the quantity, and precisely when it must be available.

Different Tools for Different Types of Variability

If there are three types of variability, this leads to a requirement for separate and distinct algorithms for planning, scheduling, and execution. The three commonly used algorithms are as follows:

- *Project Management*—Type 1 variability, which relies heavily on the concept of critical path methodology and establishing well-defined relationships of the tasks. Once the tasks have been identified, the correct sequencing will yield the project network. This network becomes the schedule for managing the resources and executing the project. Again, in [Fig. 35-5](#), the greatest uncertainty or variability is captured within the individual tasks. The project network schedule will not have any protection against variability in the resource queue. Typically, the amount of protection time for variability placed within the task is two or three times the actual productive time required.
- *Production Floor Scheduling*—Type 2 variability, which relies on developing well-defined relationships of the tasks and identifying resources. This algorithm does not use the concept of critical path methodology. As shown in [Fig. 35-6](#), Task A, B, and C and set up and set down have very little variability. This drives most of the variability into the resource queue. In fact, if one looks at the ratio between the times scheduled to accomplish all of the tasks in manufacturing of the individual product to the productive time ([Fig. 35-5](#)), which is the actual touch time needed, this confirms most of the time in the schedule is placed in the resource queue. It is not uncommon to schedule the manufacturing cycle time with 10, 20, or more times than the actual touch time (Schragenheim and Walsh, 2004).

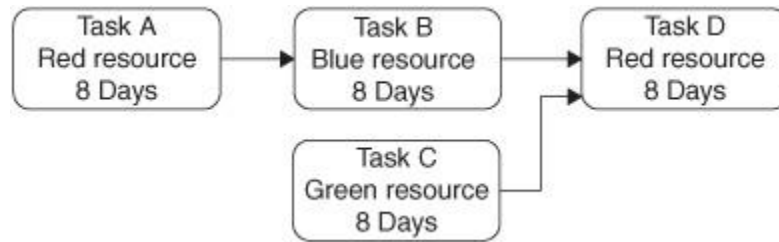


FIGURE 35-6a Traditional project network with buffering within each task.

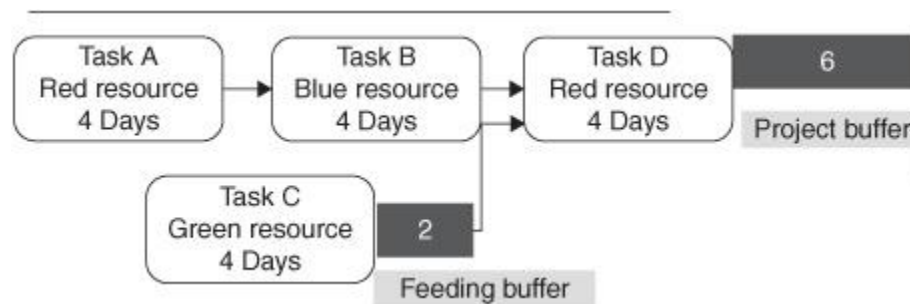


FIGURE 35-6b Critical chain project network with strategic time buffers.

- *Material Management and Inventory Control*—Type 3 variability is managed by providing safety stock of specific physical parts and finished goods (stock buffers) to protect against the changes in forecasted demand patterns. This also requires scheduling and managing material that may be outsourced or is provided directly by multiple suppliers. Material requirements have to be carefully coordinated using tools like MRP II to support the company's schedules. In addition to receiving material from suppliers in order to manufacture the company's products, this also includes scheduling materials to and through the stocking points to their final destination through the distribution channels.

Regrettably, most companies feel that, in spite of their best efforts and willingness to implement a multitude of process improvement initiatives, they fall short of achieving the anticipated returns. The reason companies fail to achieve their objectives significantly is a lack of focusing on improving the entire system as depicted in [Fig. 35-4](#). Rather, the tendency is to focus on improving individual functional areas of the company without truly understanding the net effect it will have on the profit or return on investment. [Figure 35-4](#) shows the local variability experienced by the individual functional areas of the company (departments, work centers, etc.). The source of the variability may be caused by disruptions within the functional area or other functional areas, late deliveries from suppliers, or changing market demand patterns.

Defining the System

The first step in developing a systems approach to improving the company is building a combined work flow diagram of the design, production, and distribution and supporting networks. Starting at a high level ([Fig. 35-4](#)) will lead to additionally granular diagrams until you have defined the level of required detail. A word of caution—keep the work flow diagram at a fairly high level or you will get bogged down in needless detail. The diagram can be developed further with as much detail as needed when action plans are being developed. This macro-to-micro approach has proven helpful in analyzing and creating effective company systems architecture of the types of planning, scheduling, and control systems. At times, the different types of variability may appear not to be that clear cut; if so, I encourage you to make the effort to identify which type of variability is involved. This effort will give you a better understanding of what lies ahead. Perhaps it may

be a hybrid environment where more than one algorithm must be implemented as part of a system.

The TOC Approach

Regardless of the source or cause of variability, it is far more important to know how it is impacting the company rather than just how it is impacting an individual functional area. Variability is the key indicator of how valid your assumptions are and how well the planning is being executed. In other words, if you are measuring this variability it will be an indicator of how effectively the planning and scheduling is deviating from what you thought was going to happen. However, in order to do this monitoring, there must be a common metric tying all of the individual functional areas to the company's Throughput. This metric is *time*. By using time as the overarching metric, it is now possible to evaluate if the individual functional areas throughout the company are staying within a predetermined acceptable *time burn rate*. It is now possible to see if the variability is consuming an unacceptable amount of time. This provides a process for evaluating the potential impact the variability in any part of the company will have on performance. TOC focuses on *time management* and ties this to the disruption this variability is causing within the schedule.

Visualize a time bank, referred to as a *time buffer*, providing additional time to individual functional areas if needed to protect the schedule from variability. Then the time buffers are placed strategically in the schedule, providing significant protection while protecting the delivery dates of your products or services being provided to clients. Once this connectivity is established, then it is possible to monitor the time buffers. This is called *Buffer Management* (BM). Buffer time and BM are new and important concepts first developed by Eliyahu Goldratt while developing TOC,

and are the basic building blocks for the business strategies and solutions discussed next.

TOC recognizes the existence of interdependency and variability in all organizations; in fact, all of the TOC business solutions are firmly grounded in these tenets, providing tools to better leverage the organization's Throughput. The interdependencies of the different functional area resources and their corresponding statistical fluctuations, which are manifested as variability, are shown in [Fig. 35-4](#). The three TOC business solution algorithms are as follows.

Project Management

Critical chain is the longest path recognizing task and resource dependency.⁶ Time buffers of aggregated safety are placed strategically throughout the project, providing much greater protection against variability for the critical chain than conventional critical path methodology. During project execution, monitoring the individual rate of buffer penetration against predetermined acceptable levels will provide real time risk management information. In most cases, this information will be provided early enough to allow for the required action to be taken before the promised delivery date is impacted.

[Figure 35-6a](#) is a project network where task A (using Red Resource) is scheduled to take 8 days to finish. The successor tasks when completed will feed into task D, the last task in the project. Traditional project management tools are typically used to schedule work in a Type 1 variability environment. Project management does not normally have a resource queue to provide protection against variability in the scheduling algorithm. Everyone knows that in execution, variability will cause many of the tasks to take longer than anticipated, so the common

practice is to embed additional safety time within the task itself.

The TOC project management algorithm, *Critical Chain Project Management* (CCPM), removes the protection or safety time placed in the individual tasks and schedules only the known time. Then part of the total time removed is placed in the high-risk integration points as a feeding buffer throughout the project network. An additional portion of the removed safety time is placed after the last task as a project buffer. The critical chain is task A + task B + task D and when combined with the project buffer placed at the end of the last task we establish the duration time of the project. In essence, removing the safety time previously embedded in the individual tasks and strategically placing 50 percent in time buffers throughout the project provides much better protection from variability by aggregating the safety time at strategic points (see [Fig. 35-6b](#)). This buffer protection allows for establishing control limits and monitoring the rate of time penetration into the feeding and project buffers, providing valuable real-time information of precisely when and where variability is affecting the project. This is crucial for effectively prioritizing where the resources are used when you are resource limited. A more in-depth explanation of the critical chain solution can be obtained in the book *Critical Chain* (Goldratt, 1997) and in Section III of this Handbook.



FIGURE 35-7a Serial line showing product/service flow.

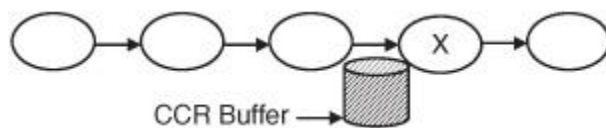


FIGURE 35-7b Serial line with a buffer inserted prior to the capacity-constrained resource.

Production Floor Scheduling

Drum-Buffer-Rope (DBR)⁷ provides buffer protection against variability at the most critical parts of the operation. Monitoring the buffer penetration will indicate when and where action must be taken, ensuring very high on-time deliveries. This scheduling algorithm is typically used in a Type 2 environment, where the task itself has low variability and there is a considerable resource queue. Therefore, the most fertile area for reducing cycle time is not in improving the time to perform the task but rather reducing the queue.

[Figure 35-7a](#) depicts a simple routing of tasks required to build a product in a manufacturing process in a Type 1 environment. The routing is built in isolation and is subsequently added to the master schedule, which is used for scheduling many other products. In conventional scheduling algorithms, the paradigm is one of loading the master schedule until every resource is fully utilized. However, the DBR approach schedules the constrained resource to no more than 85 to 90 percent (in simplified DBR), which provides a time buffer for protecting the constrained resource against variability. (See [Chapter 9](#).)

The *capacity constrained resource* (CCR)—*a resource that if not managed effectively will become the constraint*—in this case X, (see [Fig. 35-7b](#)) has less capacity than the other resources in this process. This means that the resource determines how much can be produced. Therefore, this scheduling less than constraint capacity also means that all of the other resources by definition have additional sprint (protective) capacity to respond whenever variability is causing disruptions. A CCR time buffer is placed in front of the constrained resource, which means the resources in

front can start to work and deliver their output to the CCR before they are needed. This tying the rope from the CCR to the gating operation allows delaying release of the work order to the floor until a buffer time ahead of when needed by the CCR. It is common for WIP to accumulate in front of the CCR so that when variability impacts a resource, the CCR will be protected from the disruption. Whenever the disruption is resolved, the resources use their sprint capacity to catch up until the flow is back to normal. By monitoring the control limits of the buffers, management knows when and where to take action before the effects of the disruption impact delivery dates.

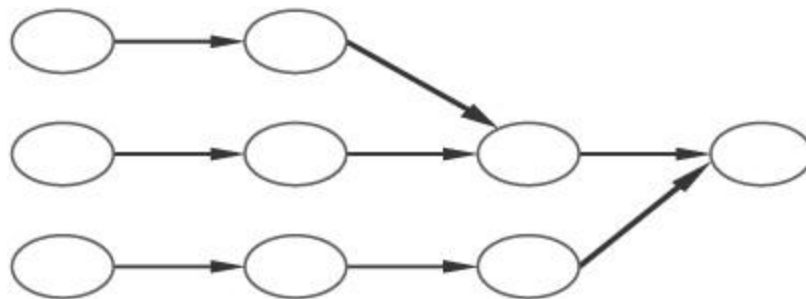


FIGURE 35-8 Typical material flow in a manufacturing operation.

This approach significantly reduces the WIP inventory, which reduces the resource queue, a prerequisite for reducing cycle time. It follows if we reduce the cycle time without hiring additional personnel, then we increase the company's Throughput.

Material Management and Inventory Control

TOC Replenishment⁸ is when stock levels are based on dynamic buffer stock levels that are much *more agile and responsive* to changing demand patterns than conventional min-max methodology.

The predominant conventional inventory control algorithm is based on determining the maximum amount of inventory carried for an item in the various stocking points in the company as depicted in [Fig. 35-8](#). These stocking points can occur anywhere needed in the production flow as required to protect Throughput. This algorithm is based on determining at what quantity level you reorder (min), triggering an order to get back to the maximum inventory level. The min level is the minimum quantity level that triggers the reordering procedure to get back to the maximum level. The min level is based on the average demand during replenishment lead time and the amount of safety stock.

The TOC replenishment approach is based on dynamic stock buffers, which are based like all TOC algorithms on managing time. This causes the inventory levels to increase or decrease in real time based on the fluctuations of market demand. Now to be clear, the stock buffers are physical material for supporting the manufacturing operations or finished product in a make-to-stock environment. The greatest source of variability is due to ever-changing market requirements for the company's products. So it may not appear that the replenishment solution is managing time; therefore, an explanation is in order. The objective is to maintain inventory levels that provide materials in a timely manner to support the manufacturing schedules. Therefore, the focus is on ensuring that as customer requirements change, the inventory levels of the needed material will be available. This change in focus, unlike the min-max methodology, allows for more frequent ordering to replenish current demands and the continuously changing trends. This allows the carried inventory in a company to be aligned closely to the market needs with significantly reduced levels of inventory. An agile and responsive material management and inventory control solution is needed for supporting the

internal critical chain and DBR schedules that are producing higher and accelerated Throughput levels of performance.

Throughput Accounting for All Methods

The TOC approach provides a common overarching metric, Throughput (T) *dollars which are defined as sales (S) dollars minus Truly Variable Costs (TVC) dollars as a rate; that is, T in dollars per period of time.* The significance is managers now have an unburdened, absolute, and real measurement that can be used across the organization. Every work center, department, and functional area has a common metric on which they can focus. This enables companies to make decisions focusing on what is best for increasing Throughput with individuals and every individual support function measured on their contribution to Throughput. As previously stated, the limiting factor to increasing Throughput is the company's constraint; therefore, it follows that all support functions must always make this metric their top priority. Now for the first time, every part of the company has the same common metric for measuring the flow of *value added* being generated. This also provides managers with the individual contribution to Throughput that each part of the company is generating.

Buffers for Time Management

The other critical contribution the TOC approach provides is the concept of *time management*. There are many very effective ways to manage operations within a business. Henry Ford used the concept of placing material on conveyor belts to control the flow. Dr. Ohno revolutionized the world of manufacturing by controlling the release of material and work performed as late as possible, thus reducing the queue, the key to improving Throughput. Dr.

Goldratt (2009) decided a more effective way was managing time; this also is a way of reducing the queue providing the advantages pioneered by Dr. Ohno. However, it also provides a means for protecting against variability by strategically placing *time buffers* that will send a signal when to release material to be worked on. In [Fig. 35-7b](#), the material is released earlier in time, a buffer time, to reach the CCR when needed as determined by the schedule. The buffer is divided into three regions: green (all is well), yellow (caution), and red (schedule is being jeopardized). The time buffers are an integral part of the TOC solution. Once the proper buffer levels are established, they also become the control limits. By monitoring the penetration of the buffers, it will indicate when and where variability is affecting the schedule allowing management to take action in a timely manner. In almost all cases, there is enough time to intervene without affecting delivery commitments. It is important to understand that buffer penetration indicates the system is experiencing disruption and monitoring and taking action when required is key to keeping the system in control.

A key factor then is the focus BM provides on the highest priority problems.

Since the three TOC scheduling and business solutions are focused on maximizing Throughput, providing risk management using strategically placed time buffers provides the basis for a powerful means of improving productivity across the enterprise. It appears that, using common metrics, we can now assess the impact any specific task is having on any part of the organization even though they may be using different scheduling algorithms and may even be in different functional areas.

Think of *time buffers* as aggregating a portion of the total required time and placing it in strategic parts of the schedule in order to provide significantly more effective protection. This is in stark contrast to conventional

approaches that simply release material much earlier than needed to allow additional time to combat variability. It also addresses the challenge, when using Dr. Ohno's approach, of having to store physical inventory throughout the manufacturing process or improving every single process. This will drive the work or availability of services being provided to commence earlier in time. The buffers are also control limits; therefore, in execution we can monitor how much the time buffers are being expended, which is simply a reflection of how much variability is impacting the schedule. If we have previously determined the acceptable level of buffer burn rate, then it is easy to observe if our system is in control or if any action must be taken. If action must be taken, we will precisely see exactly where the action must be taken.

Applications

An example of a systems approach using TOC tools for managing the development of bio medical devices is depicted in [Fig. 35-9](#).

A mid-size company developed new pharmaceutical and biomedical devices. Typically, the company takes a partially developed new product through the R&D phase, then laboratory testing, and then clinical trials. They would build a manufacturing plant to provide the product for testing. Every step in this very complex and exacting process must comply with the Federal Drug Administration (FDA) requirements and is subject to close oversight at every step. It is very expensive and can take years to obtain FDA approval.

Once FDA approval is achieved, this company would deliver the new product to one of the large multinational companies who in turn markets, mass produces, and sells the product. The benefit of reducing the new product

development cycle is very significant. The large multinational companies fund the entire development effort at great expense; it could range from tens of millions to hundreds of millions of dollars. The company that brings the new product to market first will end up owning the market and will always be the predominant supplier. This is a very high stakes game indeed.

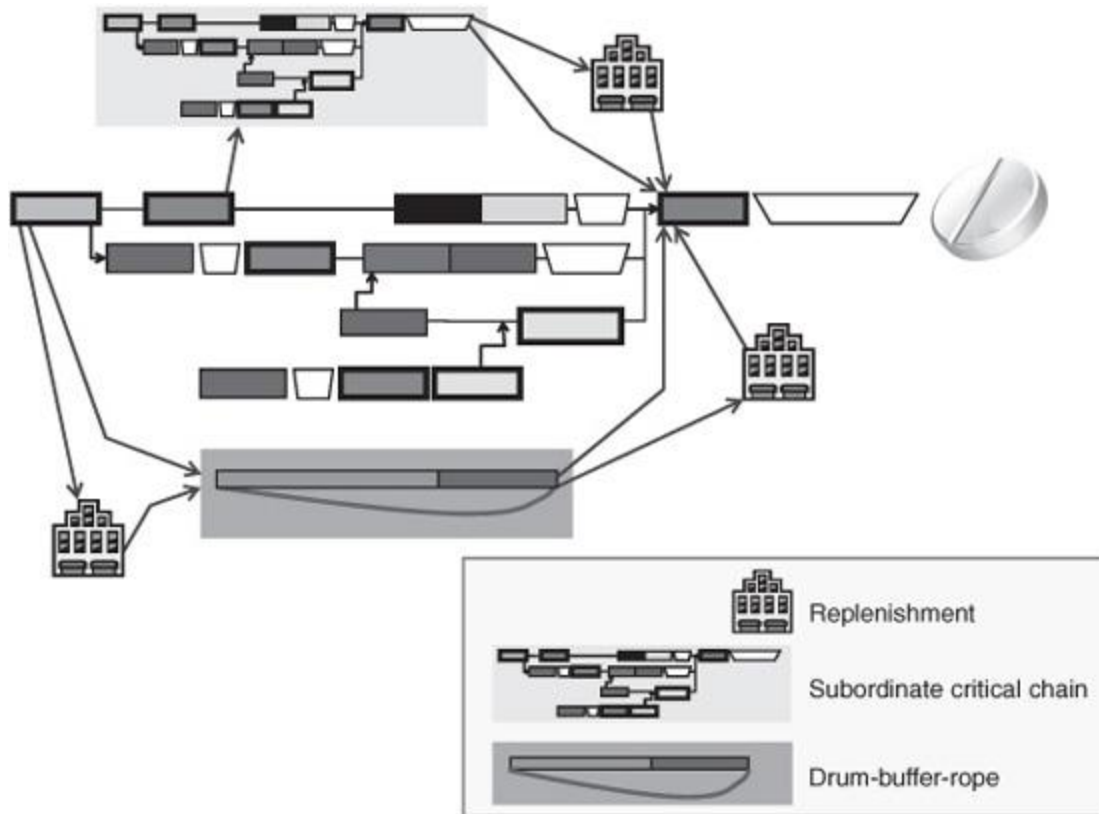


FIGURE 35-9 Integrated scheduling algorithm for a new product development.

The company has all three of the different types of variation present in their operations. As discussed earlier, it is crucial they recognize this and develop a synchronized solution set. In [Fig. 35-9](#) the overall, or master schedule, is notionally presented as a 14-task critical chain project. This is Type 1 variability, where most of the variation is within the tasks themselves. The white “tasks” are not actually

tasks, but rather are the aggregated time buffers that provide protection to the project when disturbances to the schedule happen. These buffers push the tasks to start earlier in time and since the safety time previously embedded in the tasks is removed, the duration of the project is significantly less while providing much greater protection.

The construction of the manufacturing plant is depicted in [Fig. 35-9](#) as a subordinate critical chain project, which is also Type 1. The construction of the plant is synchronized to finish and be operational when required by a task on the master critical chain schedule. The company delayed starting construction of the plant by six months and it was completed and operational with time to spare. This allowed the company, in essence, to have an additional six months before the production line had to be baselined. The maturity of the product, due to having the results of six additional months of data, was such that zero changes were made to the schedule.

When the manufacturing plant became operational, it was now subject to Type 2 variation. The manufacturing lines, the processes, and the individual tasks had very little variability as required in order to obtain FDA approval. The scheduling algorithm used was *Simplified Drum-Buffer-Rope* (S-DBR; Schragenheim and Walsh, 2004), a version of DBR developed by Eli Schragenheim, (See Chapter 9) which released the raw material for manufacturing the product, delivering to the task on the master critical chain project in a timely manner. The manufacturing time was 50 percent less than what the company historically had taken on similar products.

The TOC replenishment algorithm was used to manage the material requirements for the entire company. This requirement is a Type 3 variation, subject to rapidly changing requirements of the product development cycle. The quantity of material being held in stock was significantly

reduced, which made it easier to manage. The different sizes of the product being manufactured changed often; this is equivalent to the product mix changing often, and the replenishment solution allowed the plant to be visibly more responsive.

The next application of this approach, depicted in [Fig. 35-10](#), is at the United States Marine Corps (USMC) Maintenance Center in Albany, GA. They are one of two *maintenance repair and overhaul* (MRO) activities that service all of the USMC tracked vehicles.

The vehicles are returned to be serviced after many years of use in the field, much of which has been in very demanding environments. The mission is to return the vehicles to almost new condition, as quickly as possible and at the lowest cost. In addition, many upgrades are designed, manufactured, and concurrently installed as part of the total effort. In addition, the condition of the vehicles is unknown until inspected. Furthermore, the demand pattern is unpredictable, which only adds to the uncertainty in an already complex scheduling environment.

The vehicle itself is scheduled as depicted in [Fig. 35-10](#) as a nominal 14-task critical chain project because it experiences Type 1 variation. The actual number of tasks is much greater because it covers the major events, such as inspection, disassembly (in many cases, only the nameplate will remain on the production line), assembly, corrosion, paint, testing, etc.

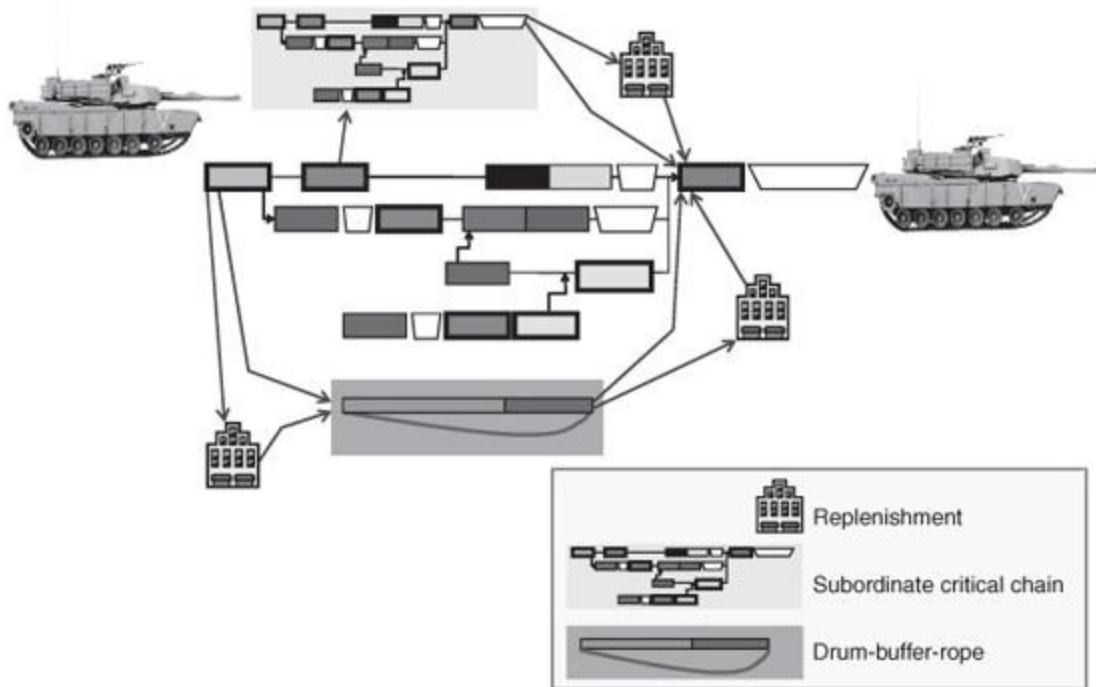


FIGURE 35-10 The integrated scheduling algorithm in an MRO environment.

Some of the major components removed for repair and overhaul such as the engine are scheduled using a subordinate critical chain schedule since again this is a Type 1 environment. The many other components that are removed from the vehicle are sent to the support shop and scheduled using DBR since this is a Type 2 environment.

The TOC replenishment algorithm is used to schedule and manage the shop consumable items and replacement parts. This is extremely challenging in a complex, high-mix, constantly changing MRO environment. This is an extremely demanding Type 3 environment of high and low product volume demand that changes on a daily basis.

All of the work has to be synchronized and come together at final assembly. This is only possible by scheduling backward from the vehicle delivery date and subordinating all efforts to the needs of the master critical chain. The uncertainty encountered in the Maintenance Center at

Albany, GA is much greater than a repetitive manufacturing environment such as when the vehicle is originally produced.

The Maintenance Center's results over the last seven years have been phenomenal. Some of the results were chronicled in the March 2005 *APICS* magazine. They were the first recipient of the U.S. Department of Defense's Robert Mason award in 2005 and again in 2007; and were the first recipient of the TOC International Certification Organization (TOCICO) Award for Excellence in 2008. They reduced the cycle time of every one of the 20 major products they support by at least half and in some instances even more. They doubled the Throughput of the organization within an 18-month period without hiring any additional personnel. They have a very ambitious process of ongoing improvement in place, which is continually raising the bar. In summary, the productivity increased dramatically, they were making their promised delivery dates, and customer satisfaction was very high.

Summary and Discussion

All of the TOC solutions use the concept of *time buffers* to provide protection against variability in execution. Therefore, for the first time we have system and operational metrics that link to each other and can be used in conjunction with the three scheduling algorithms. These horizontal and vertical linkages are crucial and provide precise information on the effect every single element is having on the company. These metrics transcend all functional areas, allowing the area managers to better understand where the real priorities lie.

This approach de facto allows for building a unified scheduling algorithm for project management, production, and distribution requirements in an organization. It solves

the longstanding dilemma of having to generate standalone schedules for individual parts of the organization vis-à-vis producing a synchronized schedule, thus providing the most benefit across the company or the entire supply chain.

First, one must have common metrics to measure how effectively each part of the enterprise is contributing to Throughput. Throughput is the rate at which each piece contributes to the output of value-added your organization generates and delivers to the market. In a for-profit organization, Throughput normally can be stated as the amount of money generated over a given period of time through sales, less the TVC. In a not-for-profit organization, Throughput could be the amount of an organization's value-added units produced per money expended over a given period of time. This approach provides the means of scheduling the many diverse functional areas and their resources in order to maximize Throughput.

In the planning and scheduling phases, all of the constraints in the organization are identified and leveraged to optimize the work flowing through the system. This pipelining is crucial and the key to producing the greatest Throughput. Therefore, this scheduling engine coordinates the projects, production, and distribution schedules by leveraging the constraints and synchronizing their efforts.

It also provides powerful and extremely effective tools for managing the inevitable variation encountered while executing the schedules. No longer restricted to continuously reacting and fire fighting, managers are given ample warning and visibility to the potential impact the variation may have on delivery dates. In most cases, the disturbance is identified quickly through BM before jeopardizing the schedules' control limits and corrective action is taken. The time buffers and BM allow management to know when they have to take action and, if so, precisely where they have to intervene. Of equal importance is now

having real-time access to information, indicating when the system or schedule is in control and no action is required.

References

- Alford, L. P. 1934. *Cost and Production Handbook*. New York: The Ronald Press Company.
- Barnes, R. M. 1980. *Motion and Time Study Design and Measurement of Work*. 7th ed. New York: John Wiley & Sons.
- Blackstone, J. H. 2008. *The APICS Dictionary*. 12th Edition, Alexandria, VA: APICS.
- Brown, M. G., Hitchcock, D. E., and Willard, M. I. 1994. *Why TQM Fails and What to Do About It*. Burr Ridge, IL: Irwin.
- Churchman, C. W. 1968. *The Systems Approach*. New York: Dell Publishing.
- Corbett, T. 1998. *Throughput Accounting*. Great Barrington, MA: North River Press.
- Goldratt, E. M. 1987. "[Chapter 5](#)—How complex are our systems?" *Theory of Constraints Journal* 1(5) New Haven, CT: Avraham Y. Goldratt Institute.
- Goldratt, E. M. 1988. "[Chapter 2](#)—Laying the foundation," *Theory of Constraints Journal* 1(2), New Haven, CT: Avraham Y. Goldratt Institute.
- Goldratt, E. M. 1990 . *What's This Thing Called Theory of Constraints?* Croton-on-Hudson, NY: North River Press.

Goldratt E. M. 1994. *It's Not Luck*. Great Barrington, MA: North River Press.

Goldratt E. M. 1997. *Critical Chain*. Great Barrington, MA: North River Press.

Goldratt, E. M. 1999. *Goldratt Satellite Program Session 2: Finance & Measurements*. Broadcast from Brummen, The Netherlands: Goldratt Satellite Program.

Goldratt, E. M. 2009. "Standing on the shoulders of giants." *The Manufacturer*. June. accessed Feb. 4, 2010 at http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants.

Goldratt, E. M. and Cox, J. 1984. *The Goal: Excellence in Manufacturing*. Croton-on-Hudson, NY: North River Press.

Hopp, W. and Spearman, M. 2000. *Factory Physics*. 2nd ed. New York: McGraw-Hill/Irwin.

Johnson, H. T. and Kaplan, R. S. 1987. *Relevance Lost: The Rise and Fall of Management Accounting*. Boston: Harvard Business School Press.

Rummler, G. A. and Brache, A. P. 1995. *Improving Performance: How to Manage the White Space on the Organization Chart*. 2nd ed. San Francisco: Jossey-Bass.

Schragenheim, E. and Walsh, D. P. 2004. "The distinction between manufacturing and multi project and the possible mix of the two," *APICS Performance Advantage*, February, 42-46.

Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>

Taylor, F. W. 1911. *Principles of Scientific Management*. New York and London: Harper & Brothers.

Zandin, K. B. and Maynard, H. B. 2001. *Maynard's Industrial Engineering Handbook*. 5th ed. New York: McGraw-Hill.

About the Author

After a successful career in leading large organizations including Director of Operations for a \$5 billion enterprise and the Executive Officer for a \$750 million aircraft overhaul and repair facility, **Daniel Walsh** founded Vector Strategies, a TOC-focused company. He and Vector Strategies are recognized experts in developing and implementing powerful strategies that quickly and dramatically improve market presence and profitability. He has worked with companies in the pharmaceutical, construction, engineering, manufacturing, aerospace, and defense industries. Numerous Fortune 100 companies are among his clients, including Textron, IBM, Caterpillar, Boeing, Lockheed, and the U.S. Department of Defense.

Daniel Walsh's success is based on his extensive experience as an executive and thought leader, as well as his development of innovative and cutting-edge systems architecture and value-added networking techniques. His focus is firmly grounded in the tenet that real and sustainable improvements in an organization must be measured on how successfully they increase profitability through value innovation.

His current efforts are focusing on developing synchronous enterprise value chain solutions in multiple

industry sectors. His research and development is centered on identifying the need to identify and leverage the strategic constraints of the enterprise, which is the key to increasing Throughput. This culminated in the development of the Integrated Enterprise Scheduling[®] (IES[®]) solution engine. Initial empirical results from deploying the IES[®] in a dozen large companies over a five-year period have been very promising. Many executives and thought leaders are convinced this may very well be the unified scheduling solution required for maximizing the profit of an enterprise-wide value chain.

Daniel Walsh has been on retainer to the Institute for Defense Analysis, a leading strategic think tank in Washington, D.C. and is a trusted advisor to many senior corporate executives. Currently he is a member of numerous corporate boards and in addition is chairman of the board of the Theory of Constraints International Certification Organization, which is dedicated to setting the standards, testing, and certifying competency in TOC.

CHAPTER 36

Combining Lean, Six Sigma, and the Theory of Constraints to Achieve Breakthrough Performance

AGI-Goldratt Institute

Introduction

As global competition continues to grow, the pressure to improve becomes more and more intense. Executives and managers face many challenges: increase sales, reduce cost, reduce inventory, accurately forecast future demand, find the next market breakthrough, and most of all survive! Although there are many ways to improve, many organizations have invested in at least one of the three most widespread methods of improvement—*Theory of Constraints* (TOC), Lean, or Six Sigma. In most cases, company experts have spent significant time mastering one of these three and spent time trying to show returns from their investment. As other methodologies came along, pressures shifted to using something else and came across as another program of the month. But for many, when the objective for all three is to improve the organization's performance, why did it come down to an "either-or" mentality? Why did some attempts at integrating the three not show the promised returns or end up being integrated in name only? Some of the reasons appear to be:

1. The methodologies were viewed as “tools in a toolbox,” where each tool was perceived as best for particular uses.
2. Expertise in all methodologies was not available, making true integration impossible.
3. An effective integration process for the three methodologies was not developed.

Our purpose is to show how to effectively integrate these methodologies, but let’s first provide a short overview of each of them.

Copyright © 2010 by Avraham Y. Goldratt Institute, a Limited Partnership.

Lean

The origin of lean manufacturing in the United States can be linked to Henry Ford (the assembly line), Fredrick Taylor (industrial engineering), and Dr. Deming (father of quality management). In Japan, these concepts were refined and honed by Taiichi Ohno, Eliji Toyoda, and Shingeo Shingo to create what is now known as the *Toyota Production System* (TPS). As shown in [Fig. 36-1](#), Taiichi Ohno once described the goal of TPS simply as to shrink the time-line from order to cash by removing non-value added waste, muda (Ohno 1988, 9).

Ohno identified seven types of waste. There are several ways to describe these “7 deadly types of waste” that occur in a system. The most common are:

1. Overproduction—producing more than the customer has ordered. Many times producing to forecast or batching to save setups can lead to over-producing.

2. Waiting—time when no value is being added to the product or service. High levels of inventory, people, parts, or information can lead to long non-value added waiting.
3. Transportation—the unnecessary movement of parts, moving multiple times, movement that does not add value. High levels of inventory, the layout of the system, and priority shifting are just a few things that can also lead to non-value added transportation.
4. Inventory—unnecessary raw material, work-in-process (WIP) or finished goods. “Stuff” we have made an investment in that the customer doesn’t currently need. Long cycle times, “just in case” thinking, and flow issues can also add to inventory issues.
5. Motion—unnecessary movement of people that does not add value. Poor workplace organization and workplace design can lead to waste in motion. These motions at times can lead to serious health and safety issues.
6. Overprocessing—adding steps or processes that don’t add value to the customer, thinking that continuing to work on something makes it a higher quality part or service. This is considered waste when the customer doesn’t require that “extra” touch.
7. Defects—work that requires rework or, even worse, work effort that needs to be scrapped. Bad processes, equipment issues, and lack of in-process control can add to the defect problem. Obviously, the more “stuff” in the system, the higher the percentage of defects.

Recently, an eighth waste has become very common and that is the waste of not tapping into human creativity.

Logically you can see how over-producing can lead to contributing to all the other waste. All wastes can be associated with any environment, not just production. Understanding and identifying waste in the system can help target improvement efforts.

The titles, “Lean Manufacturing” and later “Lean Thinking” were coined in the United States by James Womack and Daniel Jones in the 1990s to describe the Toyota Production System (TPS) (Womack and Jones, 1996). Womack and Jones introduced us to the five principles of Lean:

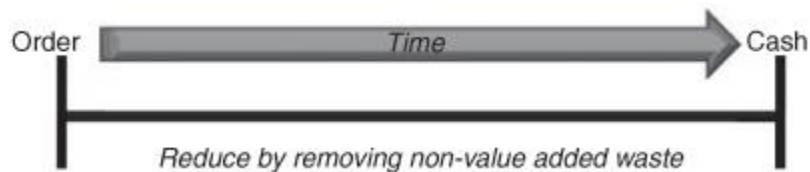


FIGURE 36-1 Goal of TPS.

1. Specify value.

As stated by Womack and Jones, “The critical starting point for lean thinking is value. Value can only be defined by the ultimate customer and it’s only meaningful when expressed in terms of a specific product (a good or a service, and often both at once), which meets the customer’s needs at a specific price at a specific time.”

The question we must always strive to answer is, “Do we truly understand value from our Customer’s Perspective—both Internal and External?”

2. Identify the steps in the value stream.

Value Stream Mapping is a process to detail and analyze the flow of material and information to bring a product or service to the customer. After identifying the entire value stream for each product, we can separate actions into value added (VA) and non-value added

(NVA) activities. Value Added activities can be defined as something that the customer would be willing to pay for; an activity that changes the form, fit, or function of the product or service and is done correctly the first time. Non-value Added is something that takes time, resources, or space and does not add value to the product, and thus adds no value to the customer. Identifying the value stream will expose many NVA activities.

3. Create smooth flow.

When the value-creating steps are understood, the next step is to create continuous flow. Things like producing in small lots versus batching, putting machines in the order of the processes, pacing production to Takt time,¹ and the application of lean tools all create smooth flow. Creating smooth flow can dramatically reduce lead time and waste.

4. Customer pulls value.

Once the first three principles are in place, we can now put a system in place that only produces at the rate of customer requirements, a “pull” system. This is the opposite of “push,” releasing work into the system based on a forecast or a schedule. No one upstream will produce a good or a service until the customer downstream is ready for it.

5. Pursue perfection.

Lean says we must continually understand value through the eyes of our customer and refine our value streams to increase the flow based on customer demands. We want to move toward perfection. The process of improvement never ends.

Six Sigma

As shown in [Fig. 36-2](#), Six Sigma has evolved from a metric, to a methodology, to a management system (Motorola University, 2008). Motorola is given credit for developing Six Sigma, but the statistical roots can be traced back to the 1800s when Carl Frederick Gauss used the normal curve for analysis and around 1924 when Walter Shewhart used control charts and made the distinction of special versus common cause variation and their link to process problems.

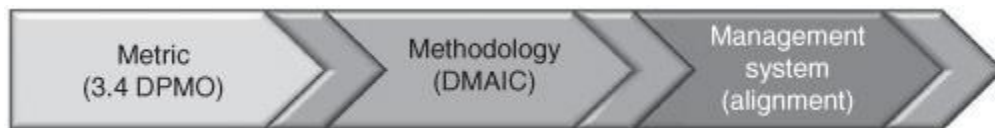


FIGURE 36-2 Six Sigma evolution.

The desired output of Six Sigma is to reduce defects, reduce cycle time, increase Throughput, and increase customer satisfaction by *reducing variation* in products and processes, thus giving an organization a competitive advantage.

Six Sigma as a metric equates to 3.4 *defects per million opportunities* (DPMO). Many companies use this metric to lead their defect reduction effort. Many improvement experts contend that most companies today work at a sigma level between 3 and 4. For example, if you are operating at a 3 sigma level, you are producing 66,800 DPMO; a 4 sigma level is 6210 DPMO. Reducing defects will obviously lead to higher customer satisfaction, lower cost of quality, increased capacity, and most important, increased profits.

Six Sigma has evolved into a business improvement methodology that focuses on how variation is affecting organizational desired results. Six Sigma project teams follow the DMAIC model to drive rapid improvement. DMAIC is an acronym for Define-Measure-Analyze-Improve-Control.

- **Define:** Typically in this stage a team is assembled, a project charter is developed, customer Critical to Quality (CTQ) requirements are defined, and a process map is created. The charter will clearly define the business case for doing the project, state the problem, define the scope, set goals, and milestones, and spell out the roles and responsibilities of team members. In identifying the CTQ issues, we must define customer characteristics that have the most impact on quality. The process map, called SIPOC (Suppliers, Inputs, Process, Outputs, Customer), defines a high-level process map of the project focus.
- **Measure:** In this step, we define *what to measure*—develop a data collection plan and perform a baseline capability study to calculate the baseline sigma.
- **Analyze:** It is important not to jump to improve before verifying why the problem exists. The main areas to look for causes of defects are data analysis, process analysis, and ultimately root cause analysis.
- **Improve:** This step takes all the data from the D, M, and A steps and develops, selects, and implements solutions that will reduce the variation in a process.
- **Control:** Sustain the new process through a robust monitoring plan.

The main purpose of the DMAIC process is for process improvement. When a process is at its “optimum” and still doesn’t meet expectations, a redesign or a new design is needed. This is called *Design For Six Sigma* (DFSS). DMADV (*Define-Measure-Analyze-Design-Verify*) is a common acronym used today for DFSS.

Motorola was one of the first companies to realize that a metrics and methodology approach was still not enough to drive “breakthrough” improvement. They continued the Six Sigma evolution into what is called the Six Sigma

Management System. *A Six Sigma Management System is a structured process to ensure that all improvement efforts are aligned to business strategy.* Six Sigma has become a top down approach to execute strategy through the alignment of all improvement activities to assure fast, sustainable growth.

Theory of Constraints (TOC)

The basic concept of TOC is often introduced through the chain analogy. A chain is only as strong as its weakest link. Improvement that does not improve the performance of the weakest link most likely does not improve the system and can be considered waste. Many claim TOC is just common sense, but it is surely not common practice.

Introduced by Eli Goldratt in the mid 1980s, a wide awareness and understanding of parts of the TOC methodology was first accomplished through people reading the book, *The Goal* (Goldratt and Cox 1984). Although many of the TOC basic concepts were discussed in *The Goal*, the complete body of knowledge was not.

Some people think of TOC as simply finding and speeding up Herbie (the fictional Boy Scout in *The Goal*), the bottleneck. Then they find the next Herbie and the next Herbie, etc. TOC is *not* about chasing Herbies. More accurately, TOC is about how to improve and manage how the system constraint (Herbie) performs in the context of the total system. This is quite different. It is about managing the total system, which is comprised of interdependencies, variability, and constraints, to ensure maximum bottom-line results for the organization. TOC is about *focusing* first on the system's leverage points and then on how all parts of the system impact the operation of the leverage points. This is the way to achieve total *system* improvement, not just localized improvements.

TOC applies the logical Thinking Processes (TP) used in the hard sciences—cause-and-effect—to understand and improve systems of all types, but particularly organizations. The process a doctor would follow if you went to him with an illness, *first Diagnosis, then Design of a treatment plan, and then Execution of the treatment plan*, is the same process followed by TOC with the use of three questions, What to Change, What to Change to, and How to Cause the Change.

One of the core beliefs of the hard sciences is that for many effects there are very few causes. Using the construct of “cause and effect” becomes increasingly important as we perform scientific analysis. All too often, we see organizations treating many “symptoms” instead of addressing the root causes. TOC looks for the core conflict that holds the root causes in place.

Think of an organization as a “money-making box” (see [Fig. 36-3](#)). It is first primed with investments in equipment and Inventory (I). Money is continually poured in as Operating Expense (OE) to pay for people and other ongoing expenses. The people process the Inventory and sell their products to make a larger amount of money called Throughput (T) (money generated by the system through sales).

The TOC systems approach requires that you first understand the system, its goal, and measurements. Then you can apply the Five Focusing Steps² (Goldratt, 1992, 307):

1. Identify the constraint(s).
2. Decide how to exploit the constraint(s).
3. Subordinate/synchronize everything else to the constraint(s).
4. If needed, elevate the system’s constraint.

5. If the constraint has been broken, go back to Step one. Do not let inertia become the constraint.

The application of these steps in a situation where the system constraint is physical is usually obvious and straightforward. However, often it is not a physical constraint. The nature of many constraints in organizations is policy constraints. In that case, the Five Focusing Steps break down into three questions (Goldratt, 1990, [Chapter 2](#)):

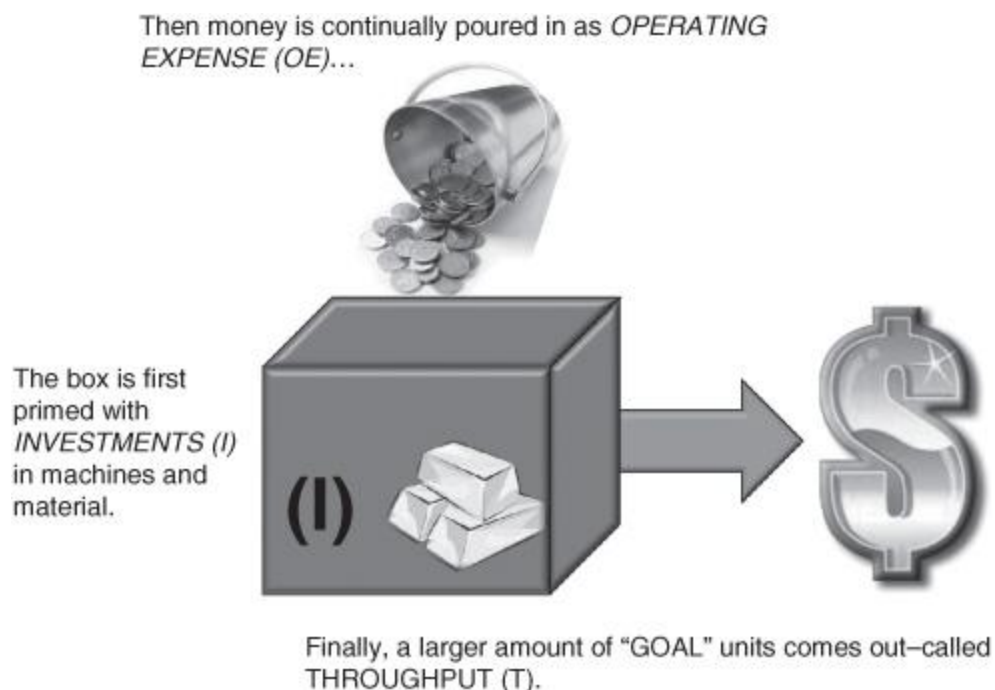


FIGURE 36-3 Money Making Box (Adapted from the “Structured Presentation” 1990, © Avraham Y. Goldratt Institute) Used with permission, Avraham Y. Goldratt Institute, a Limited Partnership.

1. What to Change?
2. What to Change to?
3. How to Cause the Change?

The TOC methodology looks at the world through the eyes of cause-and-effect logic and focuses on managing *system* constraints, interdependencies, and variability.

Discords that can Block the Effective Integration of TOC and Lean Six Sigma (LSS)

There are many synergies between the methodologies. They are all customer focused and want to provide the best value for the customer. Lean and TOC focus on creating a pull system to increase flow through the process and shorten the lead time to market. However, there are several discords between the methodologies that if not handled carefully will diminish the gains the organization can achieve from their improvement efforts.

In the early stages of the “design” of a system, there is a difference in approach between Lean and TOC.

Most Lean designs calculate Takt time, the rate at which you need to produce to meet customer demand, and then attempt to balance resources and equipment to that rate. Capacity in any operation that is greater than the amount needed to satisfy demand is considered waste. Improvement initiatives then focus on how to eliminate the waste in order to “balance” out the capacity and be equal to the demand. Due to variation, most Lean designs today will make sure that the cycle time of each operation is some percentage below Takt time, but the goal for the design of the “ideal” system is to have a balanced line with little or no “excess” or waste. In this “ideal” system, the capacity of each operation in the system would be balanced to support a cycle time just slightly shorter than the Takt time. Note that in this case, every operation in this ideal system could become the system’s constraint if there is any variation in demand, product, or processes.

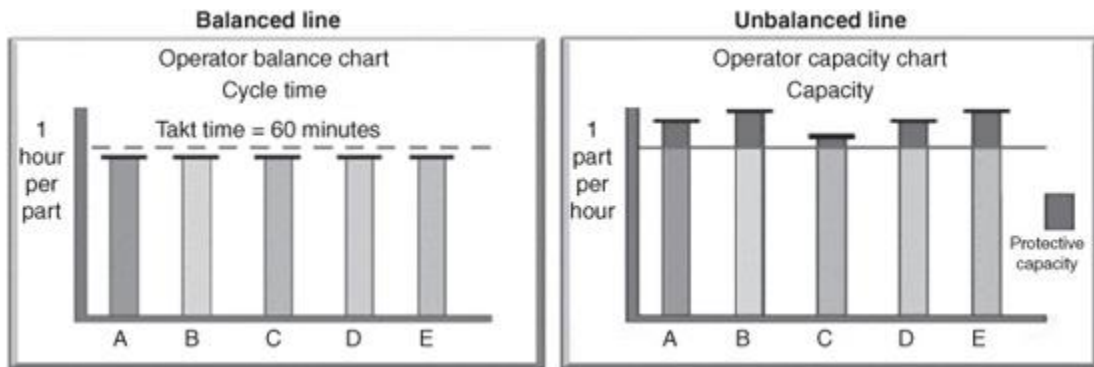


FIGURE 36-4 Balanced or unbalanced.

The TOC approach believes that there is a constraint in every system, and the constraint dictates the output of the organization. An hour lost on the constraint is an hour lost for the entire organization; thus, we don't want to "starve" the constraint. A TOC design would have some sprint or protective capacity on non-constraints to ensure that the constraint can be *exploited* to the fullest extent possible. This "unbalanced" capacity allows all operations to focus on how they are impacting the operations of the constraint and thus how their actions are impacting the Throughput of the total system. [Figure 36-4](#) shows the difference in how a balanced line and unbalanced lines are set up. When integrating TOC and Lean, the correct choice must be made.

If there is no variation, in either process times or demand, a balanced line can work. This is obviously not very likely and Dr. Deming suggests there will always be variation. An unbalanced line enables one to protect Throughput from that variation. Variation anywhere in a balanced line can immediately have a negative effect on the Throughput of the organization. Continued variation at different operations in a balanced line will dictate that you eliminate variation in the entire line in a very quick manner, which would often be a very huge and costly task. "Focus on everything, and you have not actually focused on anything" (Goldratt, 1990, 58).

The unbalanced line approach focuses on the constraint and ensures that non-constraints have enough protective capacity to catch up to the constraint when “Murphy” strikes. Eliminating variation is still a priority in an unbalanced line. The difference is that the focus of the improvements are directed to what will rapidly improve and protect Throughput while reducing Inventories (or other investments) or Operating Expenses.

In summary, both designs are set up to meet customer demand. The balanced line works well when there is little or no variation in product mix, process times, or demand. The unbalanced line works well in the presence of variation in product mix, process times, and demand. While variation reduction is a priority in both designs, the difference is where and how many places one must focus and what the impact will be on the Throughput of the total organization. The constraint in the unbalanced line is managed very tightly. Efficiency and predictability at the constraint are important metrics. The non-constraints are measured on their effectiveness in keeping the constraint supplied—this is called time buffer management. The output of the total system is the overall top metric.

Work Behaviors

The balanced-unbalanced design decision dictates how resources will be measured and ultimately how they will behave. Lines with balanced capacity expect workers to work to Takt; unbalanced lines would have workers working to the “relay runner”³ work ethic. [Figure 36-5](#) depicts the discord between working to Takt and working to the Relay Runner ethic.

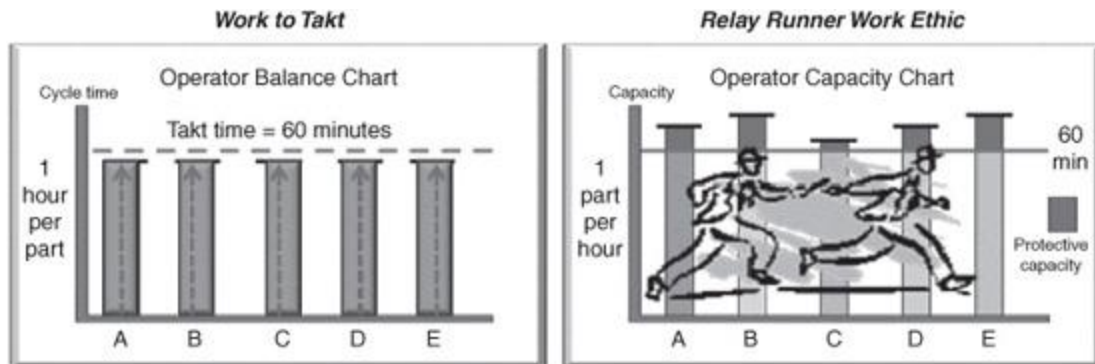


FIGURE 36-5 TAKT or Relay Runner work ethic.

Once Takt is determined and the line is balanced, the operator is to work to Takt. This works well when there is little or no variation in product mix, process times, or demand. However, if you have negative variation in the actual versus planned processing time of an operation, the work is blocked from moving to the next operation at Takt time. This results in a negative impact on Throughput and typically calls for inserting coping mechanisms on the shop floor. When there is positive variation, the worker has no incentive to pass the work on quickly so there is little opportunity to do other value added work.

Behaviors common in a work to Takt time environment are the student syndrome and Parkinson's Law. With *the student syndrome*, you think you have ample time to finish the task and therefore hold off starting the work until the last minute. If variation occurs after the last minute start, the work is finished late. Parkinson's Law states that, "Each task will expand to fill the allotted time available."

In this environment, improvements are masked due to these policies and behaviors. Early finishes of each operation are not passed on and late finishes by any operation can disrupt meeting the Takt time of the total system. This is the result of having protection that by policy is isolated within each operation and therefore cannot be aggregated to protect the total flow time. When Takt time is

violated in one operation, the entire line suffers the consequences.

The relay runner ethic emulates a finely tuned relay race team. When work is present, the operator works head down at a fast pace that is consistent with quality and safety until the work is completed or he is blocked. Should the operator become blocked, he works on the next sequenced job until the previous work becomes unblocked. This eliminates the student syndrome and Parkinson's Law effects while exposing improvement opportunities. In the relay runner environment, early finishes are passed on immediately and are aggregated to form time buffers that protect the constraint and the delivery to the customer from variation in process time or demand. Thus, the on-time delivery and Throughput of the system are protected even in the presence of significant variation.

In summary, note that how you designed your line—balanced to Takt or unbalanced—will dictate if the system works to Takt or applies the relay runner work ethic. In recent years, there have been many “workarounds” offered to try to make a balanced line work to Takt time in the presence of variation. These “workarounds” often redesign the line to an unbalanced state.

Material Release

Another subtle difference in applying TOC or LSS to a system is how material is released into a system. Both systems are pull systems based on responding to customer demand. The main difference is that the TOC signaling method is based on time, while the LSS method is based on inventory.

As shown in [Fig. 36-6](#), when there is demand on the time-based system (known as Drum-Buffer-Rope[DBR]) there is a signal sent to the constraint for scheduling purposes to

meet a shipping request, and a signal is sent from the constraint to the beginning of the line (production control) for timing the release of material. As discussed earlier, this is an unbalanced line. The non-constraint resources have “catch-up” capacity to assure orders get to the constraint on time and to the customer on time even in the presence of variation. Buffer times are calculated from the constraint to the shipping point, called the shipping time buffer, and from material release to the constraint, called the constraint time buffer. These buffers will absorb variation in getting to the constraint and to the customer, thus protecting Throughput. Material is released based on the time buffers and the actual run time of the constraint. Material is only released into the system when there is a pull from the customer; therefore, the WIP in the system is based on customer need and what the constraint can produce. There is no standard number of units of WIP, but the WIP is based on the amount of processing time that it will take on the constraint resource.

In the time-based system, high variation in demand, product mix, and process times are accommodated through adjustments to the two time buffers. These time buffers act as shock absorbers to all of the operations preceding them. Instead of providing large buffers to accommodate variation at each individual operation, the Relay Runner work ethic allows the buffering to be aggregated just in front of the constraint and in front of the customer. The protective capacity of non-constraint resources coupled with the Relay Runner work ethic allows them to catch up when there are disruptions any place in the system. Some protective capacity is usually available at the constraint resource as well. This allows it to catch up when it is the cause of disruptions.

As shown in [Figure 36-6](#), an inventory-based release system (Kanban Manufacturing System) is activated when there is a customer demand. A signal to produce, called

“kanban” is sent upstream link by link, as material is pulled to satisfy and protect customer requirements. This process is continued until all supermarkets needing replenishment are filled. The *Kanban system is a system of visual signals that triggers or controls material flow.* The Kanban in each supermarket is set to restock each part to its “Standard Level” once the signal is sent to reorder. Kanbans synchronize work processes across a system. In this system, nothing is produced unless there is a signal to produce.

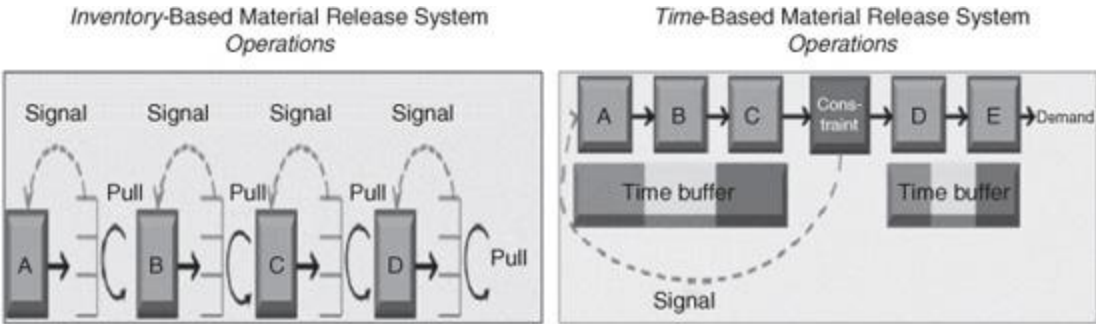


FIGURE 36-6 Release of material—time or inventory.

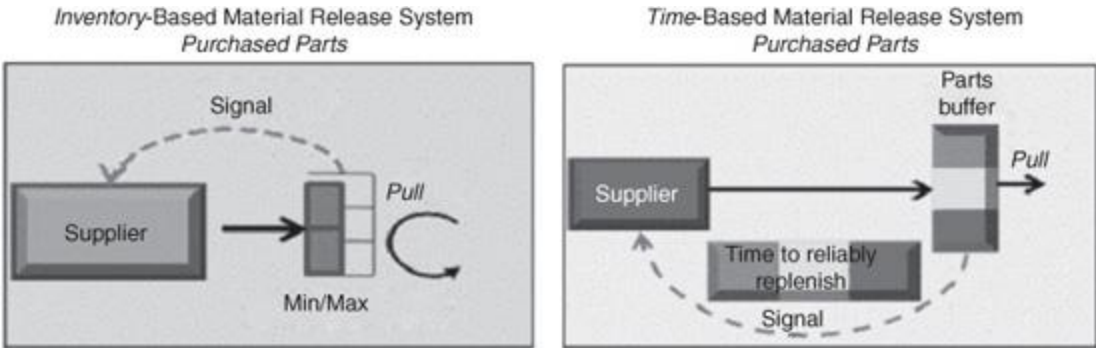


FIGURE 36-7 Replenishment system—time or inventory.

In systems with high variation in demand, product mix, or process times, the inventory-based system will not work effectively. In the inventory-based system, high variation in demand, product mix, or process times can lead to high variation in the Takt time, which can require frequent

rebalancing of a balanced line. Variation can create wandering bottlenecks, which can disrupt the flow through the line and have a negative impact on the Throughput of the system.

Replenishment System

Another subtle difference between TOC and LSS is determining the size of raw material and finished parts inventories and the difference in the mechanism for triggering the need to resupply them. [Figure 36-7](#) illustrates a traditional replenishment system⁴ and a TOC replenishment system. In a traditional replenishment system, the size of the parts inventory is based on a min-max type of system with the reorder point to resupply based on a predetermined physical quantity remaining, often known as the reorder point. TOC sizes the buffers based on demand patterns during the *time to reliably replenish* (TRR). *The TRR includes a fixed reorder time interval (e.g., once a day, once a week, etc.) and that time interval is the signal to resupply the parts inventory with what has been consumed.*

This is a *time-based* replenishment system versus an *inventory-based* replenishment system. The batch size is variable based on the demand during the fixed reorder interval. The inventory-based system has a fixed minimum batch size (the maximum level minus the reorder point) and the time interval to trigger resupply varies. The “time”-based system handles variability much better than the “inventory”-based system because the time-based system’s replenishment time is bounded. In the inventory-based system, the time to trigger the replenishment is unpredictable and can be very long.

The time-based system will work effectively in any environment. The focus is on managing the flow of parts in time versus managing levels of material. It really comes

down to what makes you pull the replenishment trigger—time or parts.

[Figure 36-8](#) reveals the design differences that you must be aware of when integrating TOC and Lean.

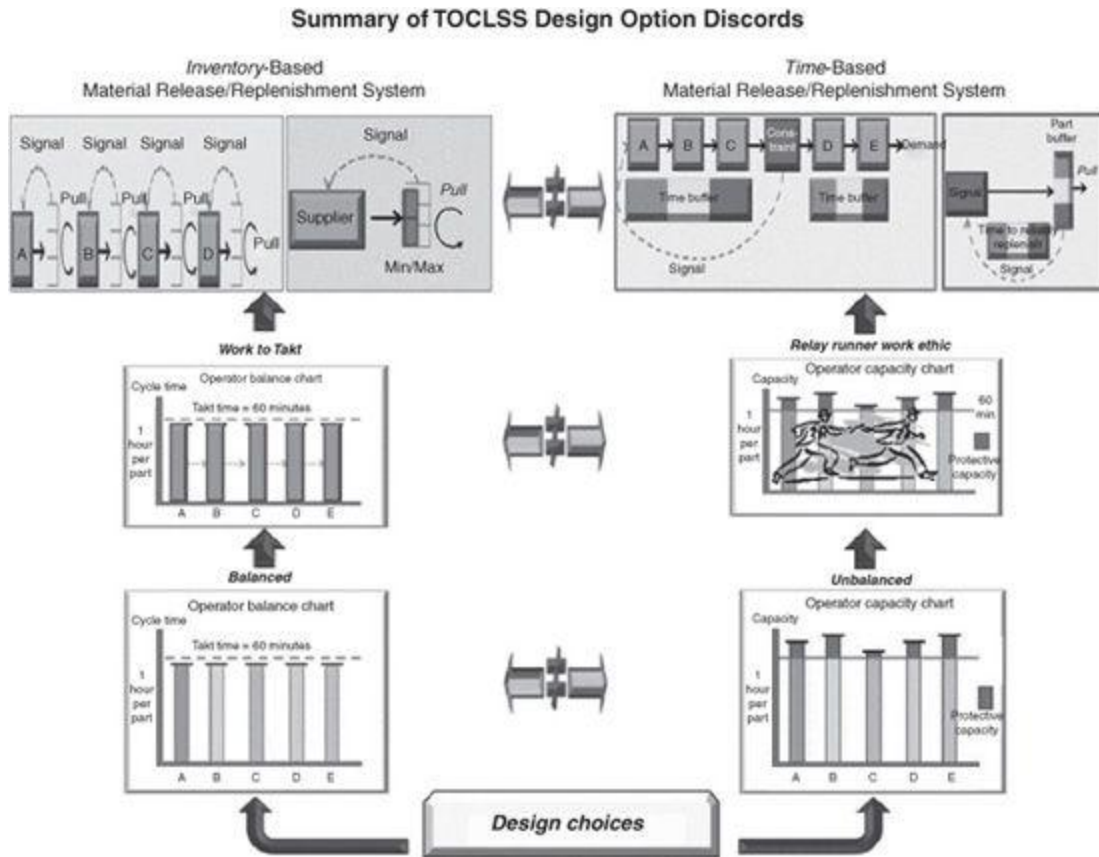


FIGURE 36-8 TOCLSS design choices.

The design choice between a balanced or unbalanced line will lead to different resource behaviors and replenishment systems. Despite what some say, the designs are *not* “the same just different”; the design intent is different and you will get different results depending on the environment. The “balanced” design works very well in the absence of demand, process time, and product mix variation. The unbalanced line, typically thought of as the best way to go in low volume high variability environments, actually works best in all environments.

How effectively we integrate the three methodologies depends on the design choice path that is taken. If the Lean design path is taken (balanced line, work to Takt, inventory release and replenishment), then only two of the TOC Five Focusing Steps can be applied—Step 1: Identify and Step 4: Elevate. These two steps will need to be applied continuously to identify and eliminate each new constraint. During this effort, the process will not be stable or in control. If the organization wants to experience the full power of the TOC Five Focusing Steps, the other design path (unbalanced line, relay runner work ethic, and time-based release and replenishment systems) must be followed. This path provides early system stability and focused system improvement.

TOCLSS—Fully Integrated TOC, Lean, and Six Sigma

The most powerful way to integrate TOC, Lean, and Six Sigma begins with strategy. The strategy provides the roadmap to improve business results. This strategic roadmap provides the direction for the areas of the organization that can most benefit the total system by applying improvements first. The system design of the first area provides predictable and stable system performance by focusing on protecting and managing the constraint(s) of the total system. Once this is achieved, process improvement efforts can be applied in a focused way to provide even more bottom line results for the organization. Finally, the improvements must be sustained in order for the organization to achieve real bottom line results over time.

In [Fig. 36-9](#), the SDAIS Model illustrates the deployment framework to ensure business success by driving effective focused process improvement through TOCLSS, from a stable operational platform.

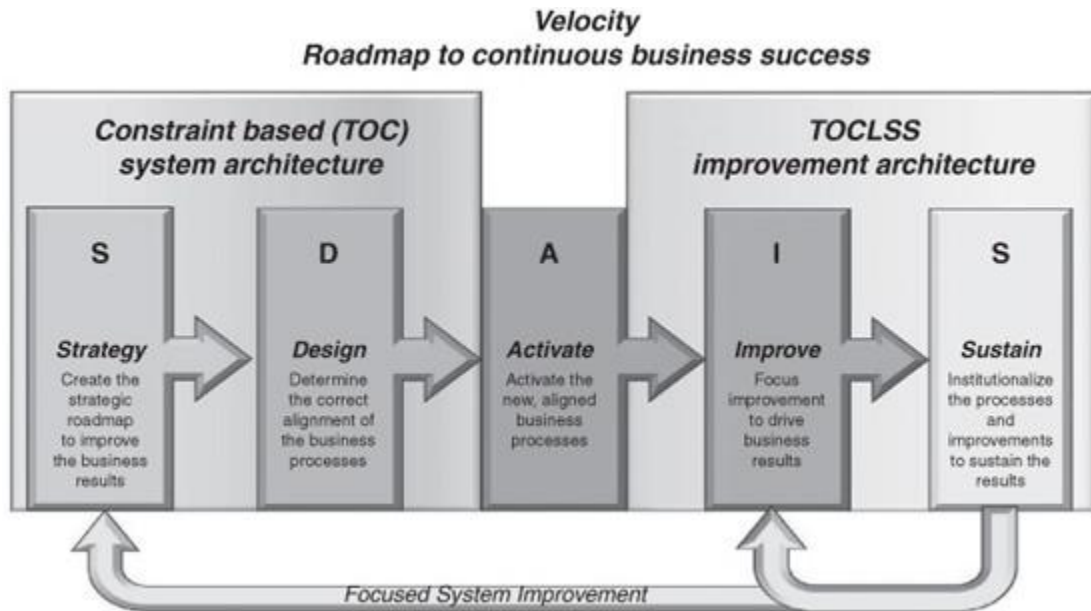


FIGURE 36-9 The Velocity Approach ((C) Avraham Y. Goldratt Institute. LP 2006-2010.)

The Velocity Roadmap to continuous business success has three major parts: the constraint-based system architecture and the TOCLSS improvement architecture, combined with the SDAIS deployment framework. In order to begin to really improve what is important, you have an understood direction and an aligned stable platform that delivers reliable, consistent Throughput.

Strategy— The output of a good strategy session is a clear, agreed upon roadmap to improve business results. The TOC strategy process involves using cause-and-effect logic to understand the core conflict of an organization, validate the conflict, and then develop the future reality, which breaks the conflict and adds other “injections” needed to improve the system. Roadblocks are removed and the result is a strategic roadmap to the future. This is done using rigorous cause-and-effect logic, which not only shows the sequence but also the interdependencies in the plan. This is much different from most strategic plans that end up being no more than an isolated list of actions from

each department. The focus is on optimizing the performance of the total system versus improving the individual departmental functions in isolation.

Design— Operational/functional leaders and subject matter experts design their operations to align their business processes to achieve the identified strategic bottom-line results. During the design process, they reconfigure the operational model, policies, measurements, roles and responsibilities, and information systems within the context of strategy and proven TOC solutions and execution management tools.

Activate— During the activation process, the newly defined policies, measurements, roles and responsibilities of the operational model and the information systems, and execution management tools are implemented to make the design operational.

This constraint-based system architecture will produce a system where business processes are designed, aligned, and operated in a stable, predictable manner.

Once a system is stabilized and is delivering stable predictable results, ongoing focused system improvements are applied that result in increased sustainable bottom-line results. TOCLSS uses the synergy of TOC, Lean, and Six Sigma to coherently achieve *focused system improvement* (FSI) beyond what might be accomplished by applying each method individually with a traditional *continuous process improvement* (CPI) approach.

Improve— Once a more stable operational system exists, the energy is turned to focused improvement efforts to drive the operational system to achieve the desired effects and strategic objectives identified in the organization's strategy session. Improvement efforts are evaluated based on their ability to increase Throughput, and to reduce Inventory and Operating Expense and advance overall system performance (Jacob, Bergland, and Cox, 2009). *Key performance indicators* (KPIs) are examined to identify gaps

between present and desired performance levels. The gaps are analyzed further and opportunities are assessed to focus improvement efforts at the business process level to achieve the desired outcomes. Improvement experts determine which improvement technique(s) are needed and then identify improvement project priorities. Some useful improvement techniques include 5S System, Standard Work, Rapid Setup Reduction (SMED), elimination of non-value added waste, Total Productive Maintenance (TPM), Point of Use Storage (POUS), Mistake Proofing (Poke Yoka), Visual Tactics, Control Charts (SPC), Capability Studies, and Design of Experiments.

Sustain— Organizational memory is created and supported by establishing the documentation of the strategy, operational design and the focused system improvements details. The organization continually reviews key measurement results to assess, address, and institutionalize the policies, measurements, and behaviors to guarantee that the results are sustained and do not degrade. The organization ensures that they have continued capability to achieve buy-in and maintain expertise.

Following the SDAIS process eliminates the need for an organization to have to “choose” a methodology, or to use the “toolbox” approach randomly. The organization can utilize the full integration of TOC, Lean, and Six Sigma in order to obtain focused system improvement that achieves real, sustainable breakthrough performance.

References

Blackstone, J. H. Jr. 2008. *APICS Dictionary*. 12 ed. Alexandria, VA: APICS.

Goldratt, E. M. 1990. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Great Barrington, MA: The North River Press.

Goldratt, E. M. 1990. *What is this Thing Called Theory of Constraints and How should it be Implemented?* Croton-on-Hudson, NY: North River Press, Inc.

Goldratt, E. M. and Cox, J. 1984. *The Goal*. Great Barrington, MA: The North River Press.

Goldratt, E. M. and Cox, J. 1992. *The Goal: A Process of Ongoing Improvement*. 2nd revised edition. Croton-on-Hudson, NY: North River Press, Inc.

Jacob, D., Bergland, S., and Cox, J. 2009. *VELOCITY: Combining Lean, Six Sigma and the Theory of Constraints to Achieve Breakthrough Performance*. New York: Free Press.

Motorola University. 2008. *Six Sigma through the Years*. <http://www.motorola.com/content>. Ohno, T. 1988. *Toyota Production System: Beyond Large-Scale Production*. New York: Productivity Press.

Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>

Womack, J. P. and Jones, D. T. 1996. *Lean Thinking*. New York: Free Press.

About the Author

Since 1986, **AGI-Goldratt Institute** has enabled organizations to better align the way they operate with what they are trying to achieve—strategic bottom-line results.

AGI is the birthplace of constraint-based techniques and solutions for business success.

Many organizations and consultants trace their roots back to AGI not only for TOC, but also for how TOC integrates with other improvement methods.

AGI provides its clients with rapid, bottom-line results with what it calls *VELOCITY*—a powerful business approach combining speed with direction. **VELOCITY** consists of three pillars: TOC, the system architecture; TOCLSS, the focused improvement process; and SDAIS, the deployment framework.

SDAIS (Strategy-Design-Activate-Improve-Sustain) begins with creating and then executing the strategic roadmap to ensure business processes are designed and aligned to achieve the **strategy**. Once **designed**, the business processes are **activated** to allow the organization to operate in a stable, predictable manner with less investment and organizational churn.

Once stable, focused system **improvements** are applied to increase sustainable bottom line results. Execution Management tools and transfer of knowledge enable each aspect of SDAIS and serve as the foundation for self-sufficiency and **sustainment**.

AGI has expertise in TOC, TOCLSS, and SDAIS, with years of experience adapting each of these elements to meet the unique needs of its clients, regardless of size or industry.

AGI excels at leading organizations through successful business transformations by providing business assessment, implementation support, execution management tools, training, and mentoring.

We are motivated by making the complex manageable and enabling our clients' self-sustaining success.

CHAPTER 37

Using TOC in Complex Systems

John Covington

Introduction

The purpose of this chapter is to give the reader some ideas on how to use TOC thinking to address and improve performance of complex organizational systems.

What is a complex system? Complexity is in the eyes of who is looking at the issue and their perception. What appears to be complex to one person might appear simple to another. In order to be an effective problem solver you must be able reduce any system down to its simplest components, which may mean redefining the system. I spent a lot of my industrial career working in chemical plants. There are thousands of issues in a large continuous flow process facility—computer controls, raw material variability, operator training, sludge buildup inside a heat exchanger, hundreds of control valves, wear and tear on equipment, scheduling of rail cars, EPA regulations, etc. There are many details and tons of data. It can appear very complex. How does one ever get their arms around it all?

All systems transform something from one state of being to another. For example, you may have a chemical plant that converts air, sulfur, and water into sulfuric acid. The simplest definition of the system might be a box where sulfur, air, and water are going into the box and coming out of the box is sulfuric acid and byproducts. Perhaps you can

look at a university as one where people are transformed from one state of knowledge to another. One can then begin to add detail sufficient to describe the system adequately so it is suitable for solving. What are the dependencies and their sequencing necessary to achieve the purpose of the system? Those questions must be answered before you attempt to find a solution.

I have been using *Theory of Constraint* (TOC) concepts to solve problems since the early 1970s. I did not call it the Theory of Constraints then, I called it a material and energy balance. I was educated as a chemical engineer and early on, our professors instructed us to:

1. Define your system.
2. Determine the bottleneck.

Copyright © 2010 by John Covington.

Not much has changed from then to now, and essentially that is how one should attack a complex system—define the system and determine the bottleneck.

Again, all systems, whether complex or not, transform something from one state to another. Perhaps the best way to explain this topic of providing solutions to complex system problems is through three examples of complex organizations:

1. A conglomerate that transforms steel rods into “sucker rods” for the oil industry.¹ In this example, we will redefine the system, find the current logistical constraint of the new, better-defined system, and then address the mindset that would be an obstacle moving forward.

2. A company that makes the components of front wheel drive shafts in three different plants then assembles them in a fourth plant. All plants are scheduled by their competing customers. This case illustrates how important it is in complex systems to define the system properly. Get this step wrong and, at best, you have added a lot of time to reaching a solution, and at worst, you never address the real issues.
3. An organization that converts Non-Disciples into Disciples. This case provides a different view of Throughput being a nonprofit service provided or an intangible good. If a TOC mindset can address an issue as intangible as a Disciple, it can address anything.

These three cases are used to illustrate the complexity of the organizational environment and the simplicity of the solutions needed to achieve success. After the cases, I provide a summary listing the major insights gained from my experiences in these and other complex environments.

We Need More Sucker Rods!

Introduction

In 2007, my good friend Jeff Bust became president of the Energy Group (EG) of Dover Corporation. Dover is a conglomerate with over \$7 billion in sales, with the Energy Group making up about \$500 million. In Jeff's group, there were two companies that made sucker rods: Norris Rod, located in Tulsa, Oklahoma and Alberta Oil and Tool (AOT), located in Edmonton, Alberta.

A brief discussion of the culture within many conglomerates is appropriate. Conglomerates buy and sell

companies and have them under one big umbrella. Many conglomerates want to preserve an individual company's identity as they feel that independence causes them to perform better. The down side of this is that when you have two companies that produce the same thing and they are both being measured by their own profit and loss statement, then there is an opportunity for competition rather than collaboration. This was the case at Dover as the two companies within EG were struggling to keep up with market growth. Jeff felt Norris and AOT were working on the wrong things, were having quality problems, and there was minimum collaboration between the two companies.

Jeff needed more sucker rods and he needed them fast. There was no time to invest a lot of additional capital in equipment because EG needed to take advantage of the market when it was in an up cycle. EG also did not want to commit new capital to an old process of producing sucker rods.

What is a sucker rod? Most of you have seen an oil well that looks like a giant horse head moving up and down. Attached to the horse head is a rod that goes down inside a casing that may go 2000 feet or more into the earth. Each rod is about 22 feet long, so you need a boatload of them to reach some oil thousands of feet down. At the end of the rod, a device captures the oil and starts it on its journey to the surface where it is collected and sold for a small fortune.

Sucker rods come in various diameters, strengths, and lengths and the companies had nearly 100 stock keeping units of different sucker rods.

There are five basic steps to produce the rods:

1. Straighten the rods from the steel mill.
2. Go through a forging operation where the ends of the rods are formed.

3. Heat treat.
4. Machine/thread.
5. Paint.

At Chesapeake Consulting, we have not found too many complex systems that we could not make simple just by using the *five focusing steps* (5FS); however, we have added what we consider to be two important prerequisites:

- Define the system and its purpose.
- Decide how to measure it.

For this assignment, we were given the mission to:

- Get more total Throughput in order to take advantage of market demand.
- Develop a more unified strategy.
- Create some synergy and collaboration between AOT and Norris.

Therefore, the system on which we worked was the combined operations of AOT and Norris. Although the two plants are physically separated by 1000 miles, we looked at them as if they were one facility under the same roof.

Some History and What We Learned

Both Norris and AOT had done an excellent job of maintaining their corporate cultures, perhaps to a fault, and both were good performers. Dover has five major financial and operational criteria on which they judge companies and both companies were exceeding the goal on four of the five.

The market was good, so both companies were showing decent profits, and executives were getting nice bonuses. In addition, the longevity of the individual company presidents was much longer than the longevity of a division president. These two companies had survived many people that held the position of my friend Jeff. The major point here is that there was not a lot of incentive for either company to change. There was also a history of union friction at Norris and intra-corporation competition.

In [Fig. 37-1](#), the general process flow was raw materials to straightening, to forging to heat treat, to machining, to paint, to finished goods, and then shipping to customers.

The physical constraint of the individual plants and for the entire system was the heat-treat operations. The maximum Throughput that the EG could produce was the total of the heat-treat outputs of Norris and AOT. For a variety of reasons, heat treat was the logical place to have the internal physical constraint, so we made no effort to relocate it. Heat treat was the highest capital investment, it was relatively easy to buffer, and it was the process step where the company felt they added the most value; a step they would not consider outsourcing.

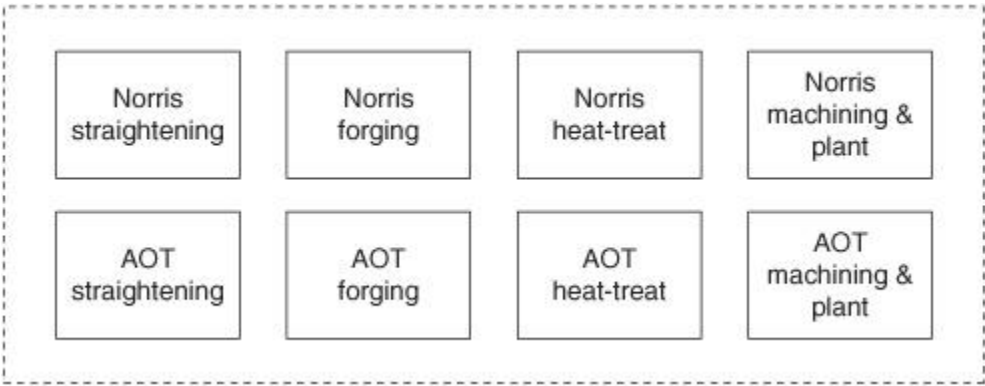


FIGURE 37-1 Combined operations of Norris/AOT.

At AOT, they had nearly balanced capacity; therefore, there was inadequate protective capacity in forging and

straightening to keep heat-treat supplied at full capacity. AOT had already applied Lean and Six Sigma tool sets and there was very little opportunity for additional capacity without making a capital investment. AOT had been working on a process of ongoing improvement for several years and were well down the path to world-class performance (whatever that means). They were using statistical methods to determine when dies needed changing and a forging changeover took less than an hour, whereas at Norris it might take over a day. There was not a lot of low-hanging fruit at AOT. In addition, a friendly culture existed at AOT; workers would smile and greet their supervisors and it was obvious workers and managers were engaged in their work.

Norris had plenty of forging equipment capacity but most of it was being wasted through sloppy operations. Although they had begun a Lean and Six Sigma journey, they had not scratched the surface and they had not focused their efforts. TOC clearly indicated that improvement efforts needed to be focused at the Norris forging process in order to gain the protective capacity to supply the combined heat-treat operations of Norris and AOT. Norris had a tendency to have long runs of rods through forging to avoid long setups. Of course, this traditional philosophy of long runs, a focus on efficiency and minimization of setup costs led to higher than desired inventory levels and consumed valuable protective capacity to make the wrong stuff.

In order to have the proper buffer in front of heat-treat operations at AOT and Norris, the system had to get more protective capacity out of forging at Norris.

The potential impact from implementing TOC was tens of millions of dollars in additional sales and reduced Operating Expense without any additional capital.

Let's pause for a second. We have taken a complex system composed of two companies under the same organizational umbrella with many culture and market issues and narrowed our focus down to one department—

the forging operation at Norris. We will start there knowing that what we do will spread to the rest of the organization.

What Change was Needed

Through an assessment, we found some of the *undesirable effects* (UDEs) were the following:

1. EG is losing sales because of extended lead times. This also had a long-term negative effect as the company wanted to remain being “the first one called.” In this industry, the phone rings and someone needs sucker rods right now or in a short lead time. The sooner a pumping station is on line, the sooner money starts flowing to the oil company. If EG cannot satisfy the customer’s order, then the customer calls the next supplier on its list and EG loses market share.
2. High scrap rate exists.
3. Hostile work environment exists at Norris.
4. Opportunities are missed to grow market share.
5. High inventory (finished goods and raw materials) exists at Norris.

The core physical issue was the forging operation at Norris. If we could wave a magic wand and make Norris forging look and perform like AOT forging, then sales would instantly increase and some other good things might happen also.

How to Cause the Change

Our approach to implementations is to put the customer through a series of experiences intended to teach them the

knowledge they need and then lead them in applying that knowledge to their specific environment. This is normally a five-step process:

1. Assessment. A good “Jonah”² had better know the answer to the question prior to asking it. During the assessment phase, I want to get a feel for what the issues are and what a solution might look like. I also want to have an idea of system definition and who needs to be involved.
2. Education. The purpose of this phase is to transfer appropriate knowledge to the client that they will need to improve. We use hands-on games and lectures laced with examples to help the customer learn in an environment that is not their own (helps avoid the “this will not work here syndrome”).
3. Design. In this phase, we get the client team to use their newfound knowledge to design a new system (complete with new policies) to improve their performance. Such a system might involve writing detailed procedures for a DBR system tailored for their environment.
4. Planning. In this phase, specific goals and tasks are defined and obstacles that would stand in the way of completion are identified. Action plans are developed to overcome obstacles. This work is similar to building a *prerequisite tree* (PRT).
5. Execution. This is where folks go out and execute the plan and progress is monitored.

It has been my experience that by the time we get through this process, most of the resistance to change has been overcome and the client has ownership of the solution.

However, in this case we had some additional reservations. As mentioned before, Norris had a history of adversarial relations between management and the union and there was an awful lot of distrust. There was also distrust between Norris and AOT. We decided that our normal approach would not work because information flows through relationships and those relationships were clogged with numerous erroneous negative assumptions. So, while the physical constraint was heat treat, the real constraint that made it a complex system was mindsets/relationships within and across plants. We needed somehow to add learning experiences that would suffice to break through the distrust issue.

What We Did to Implement the Change

Relationships, purpose (includes processes that achieve purpose), and information flow form culture, and the existing culture at Norris needed a change.

We considered the most important part of this particular implementation to be the selection, nurturing, and guiding of the implementation team. We also figured we had one shot to get it right.

To get a cross-functional team of those directly involved in forging, we figuratively went to the forges and walked outward, touching those we met. We met:

- Operators
- Mechanics
- Production foremen
- Maintenance foremen
- Process improvement engineers
- Schedulers

With a list of possibilities of team members, we sat down with upper management and human resources and selected a Norris implementation team of 15 people. Things we considered were:

- Formal and informal leaders
- Union representation
- Folks from each shift
- People with positive attitudes

Once the team was selected, education began. Process and technical education included a typical 2-day synchronous flow workshop and a 1-day hands-on workshop on Lean. The knowledge we hoped to transfer included:

- The concept of constraints
- The issue of protective capacity
- The relationship among capacity, dependency, variability, and inventory
- The strategic location of a constraint
- Ideas and tools (such as setup reduction) to gain “cheap and free” capacity
- TOC measures, such as Throughput, Inventory, and Operating Expense
- Other appropriate operations measures.

The leadership/relationship education and experience included:

- Individual assessments of behavior, values, and skills relative to projects, systems, and people
- A one-on-one session with an executive coach to review the assessments
- A 2-day communications workshop

The knowledge we hoped to transfer in this 2-day workshop was:

- Each individual of the team to better understand themselves and how their style affected others. This knowledge gives the individual the choice to adapt their behavior in order to improve information flow (remember that culture thing).
- Get to know their fellow workers/managers better to begin to build relationships based on respect and trust.

The team was now ready to head for Canada.

“Oh Canada”

Can you imagine the logistics of taking a team of workers to Canada, many of whom had never left Oklahoma? Just getting passports in a timely manner was an ordeal. I would like to pause here for a second and have us ponder the support given this project by Dover management. Some executives may have rationalized that they could have gotten the same results without the expense of sending the team to Canada. Dover executives had the wisdom to realize that this was a significant event that would send a powerful message to both companies.

AOT did a wonderful job of hosting the visit. In addition to plant tours and briefings, the joint Canadian and U.S. team had lots of food and entertainment together. The Norris team observed and participated in forging setups and other operations events. The visit accomplished:

1. A deeper appreciation by the Norris team for AOT accomplishments with respect to operational

excellence.

2. Clarity about the learning gap between facilities, which is a tangible opportunity for Norris.
3. An extensive list of team-generated ideas for both Norris and AOT.
4. Multiple new relationships among counterparts at the facilities.

Results after Six Months

Once the team returned to Norris, they implemented several of the concepts and tools learned. A summary of the important results follows:

- Increased focus on heat treat as the strategic constraint for both companies.
- Norris increased profits by 6 percent despite taking 2 of the 6 furnaces out of service for repair.
- AOT profits increased more than 6 percent and both of these improvements were achieved in spite of increased steel prices that were not passed on.
- Forging at Norris increased flexibility by reducing setup times from an average of half a shift to 30 minutes (12.5 percent of the original setup time).
- Reduced rework from a 50 percent rework rate to a 10 percent rework rate in several product lines (thus gaining additional constraint capacity for free).
- AOT now shares formulations with Norris to improve efficiency.
- More overall collaboration between the two companies.

As the implementation progressed, other changes were made to put the two companies more in alignment. One of the major changes was putting the CEO of AOT over both companies. Another was a change in the culture (mindsets/relationships) within and across plants.

Have You Really Defined the System?

Introduction

GKN Automotive is a company that produces half-shafts for front-wheel-drive cars. Their customers included Ford, General Motors, Toyota, Honda, and most other major automobile companies. Their president and CEO, Tom Stone, felt that he could improve performance by embracing the principles of TOC.

What is a half-shaft? If you climb up under your front-wheel-drive vehicle, you will note two small shafts that connect the transmission to each wheel. At the end of each shaft is a forged piece of metal that looks like a tulip—in fact that is what GKN folks call it. The end that attaches to the transmission is the “inboard end” and the one that attaches to the wheel is the “outboard end.”

There were four plants in the system. One plant produced the forged “tulips,” one plant machined the tulips and inners for the inboard side, another plant did the same for the outboard side, and both of these plants shipped to the assembly plant which assembled product for approximately 22 different car models.

All four plants were structured as separate cost centers and the entity of concern (or system) was the physical plant. During the assessment phase, it was clear to us that this structure made no sense. Looking at the GKN system from

the perspective of four independent plants measured by cost actually made GKN more complex. What we learned is that all of the machining equipment was designated for a particular car model and that the auto manufacturer controlled “their” machines. It became apparent based on the outside control of scheduling of the plants and the dependent nature of the parts and processes across plants that considering each physical plant as the system was erroneous. Each plant had to deal with parts from all 22 car models and each plant’s management was focused on “local optimization” of their own plant. (See [Fig. 37-2.](#))

If the system is not defined properly, then identifying a constraint or doing a *Thinking Processes* (TP) analysis is meaningless. In addition, it should always be our goal to look at a complex system in a manner that makes it less complex.

What Do We Need To Change?

GKN had been working on the wrong system—the physical plant location and treating them as independent cost centers. The individual plants saw themselves as having their own customers, all of which had different requirements and tendencies with respect to lead time, quality, and other issues.

What Do We Change To?

What did make sense was to segment GKN by specific models/markets. Now team members from the forging, inboard, outboard, and assembly facilities all worked to satisfy a specific customer and model. The focus became the customer (from the beginning of the process to the end) with fluid communications through the organization. This new view of the organization focused on value to the

customer. Value lanes by customer was a system that made sense. We design a value lane by starting at the customer and working back toward the receiving dock and stopping at a natural point of divergence (see [Fig. 37-3](#)). Each value lane had its own assembly, inboard machining, outboard machining, and its own special forged tulips. Operators assigned to these machines were full time and they considered themselves more a part of their value lane than a particular plant location.

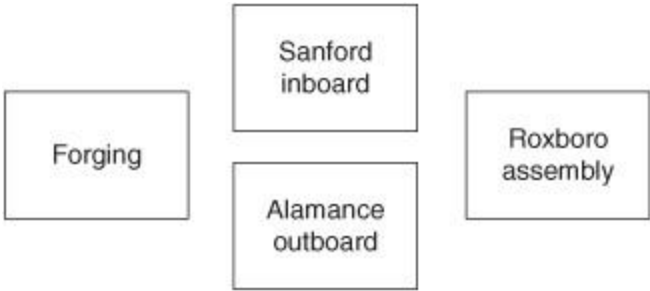


FIGURE 37-2 GKN “plant perspective.”

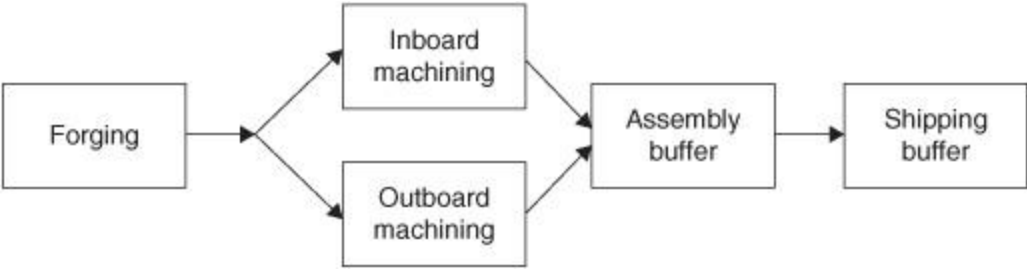


FIGURE 37-3 “Value lane” perspective—12 to 14 value lanes covering 22 models.

How Do We Cause the Change?

In the case of GKN, the concept of value lanes replacing the physical plants as the “system” was sold to the President and CEO; therefore, there was no debate as to whether we

were going to try to go in that direction. I chose those words carefully—there was no debate that we were going to try. Whether it was going to work was going to be totally dependent on the buy-in of those involved.

By changing to value lanes, we essentially began to change the functional organizational structure. In a major complex system, this change is not uncommon. The new organization consisted of 12 to 14 value lanes that covered all 22 models and each lane included workers in all four plants. The lanes were responsible for ensuring customer satisfaction for their customer—Ford Taurus, Toyota Camry, etc. Telephone, e-mails, and buffer data were modes of communication. Many hourly employees were now involved directly on what was to be produced and how things were managed. Each value lane monitored their shipping buffer and controlled release of raw materials to the processing equipment. In the transition from a functional organization structure to a DBR flow structure, plant managers and other supervisors needed to let go of control and trust their workers more.

Results

GKN experienced all of the normal TOC success:

- Net profit nearly doubled (increased by 85 percent).
- Inventory decreased by 22 percent.
- Return on net assets increased by 40 percent.
- Value added per employee (rough estimate of T/OE) increased by 30 percent.

Their executives explained their new operating method as bringing “calm to a chaotic environment.”

Where is the Constraint in Disciple Making?

Introduction

The United Methodist Church (UMC) is the second largest Protestant denomination in the world. Its founder, John Wesley, is credited by many as helping England avoid a revolution similar to the French Revolution. John Wesley preached to the poor and was all about Christians being in action to do good. He had enough of an impact that the hope he gave people through God staved off a violent uprising of the masses.

The UMC has been in decline for over five decades and they are losing about 1000 members per week. This number would be worse if it were not for the fact they are growing overseas, especially in Africa.

In the early 1990s, I was invited to do a one-day TOC seminar in Nashville, Tennessee for their General Board of Discipleship, which is one of the church's most influential agencies. Ezra Earl Jones was the General Secretary in charge of the Board and his position was on a level with a Bishop. Ezra Earl was considered an innovator in the church and attempted to use many tools that were used successfully by industry in order to turn around the plight of the UMC.

The 1-day workshop contained the TOC basics of measures and the 5FS. This session went well and Ezra Earl and his publisher of the Upper Room devotional subsequently attended an open audience Jonah Class I conducted at Clemson University. After years of working with the UMC, I now have grown to appreciate what a commitment Ezra Earl made by spending two weeks away from his job.

After the Jonah 2-week workshop, Ezra Earl said that his next goal would be to introduce the TOC concepts to several

Bishops and other leaders in the church. The opportunity presented itself in January 1996 when I spent four days at a retreat with several Bishops and other leaders at “The Grove” near Ashville, NC, which is the mountain retreat owned by the Billy Graham organization. There were several speakers for the retreat, Dr. Margaret Wheatley, author of *Leadership and the New Science*, spoke one day, Peter Block, famous author and consultant, spoke one day, I spoke one day and then I facilitated the last day where we tried to pull the three days of knowledge transfer together. It was a great session; it motivated some of the leaders to consider more study and analysis.

One thing you are going to learn in this example is that some complex systems take years and perhaps decades to begin to shift their thinking and behavior. One key lesson is not to give up. Many times an outside disrupter does not control when a shift in thinking occurs. The best we can do for the complex system is to be persistent and ready.

After this program, Ezra Earl said that his next goal would be to get a team of Bishops and other church leaders to dedicate time to do a complete analysis of the UMC and to develop a solution. He accomplished his goal and Chesapeake facilitated the workshop. Lisa Scheinkopf, who worked for Chesapeake at the time, and I would conduct the sessions. We met 12 days over several months with sessions in Atlanta, Chicago, and Nashville. While in Chicago, we stayed at a Catholic convent. Lisa, who is Jewish, said that she was way outside her comfort zone sleeping under a cross. As you might imagine, we had a lot of laughs.

The Analysis

Here are some of the UDEs the team surfaced during their TP analysis that blocked them from improving:

1. UMC lacked clarity of purpose/vision.
2. UMC is doing a poor job in spiritual formation (making Disciples).
3. Generally, the church is doing a poor job of transforming people.
4. There is no process for preparing leaders.
5. The UMC lacks a climate for innovation.
6. Spiritual malaise is prevalent throughout the church.

Lisa and I facilitated the UMC leadership through a full TP analysis with the core problem being that spiritual leaders were not fulfilling their specific roles as spiritual leaders. We came to that core issue from two directions: the TP analysis and a simpler analysis of the process required to achieve their purpose.

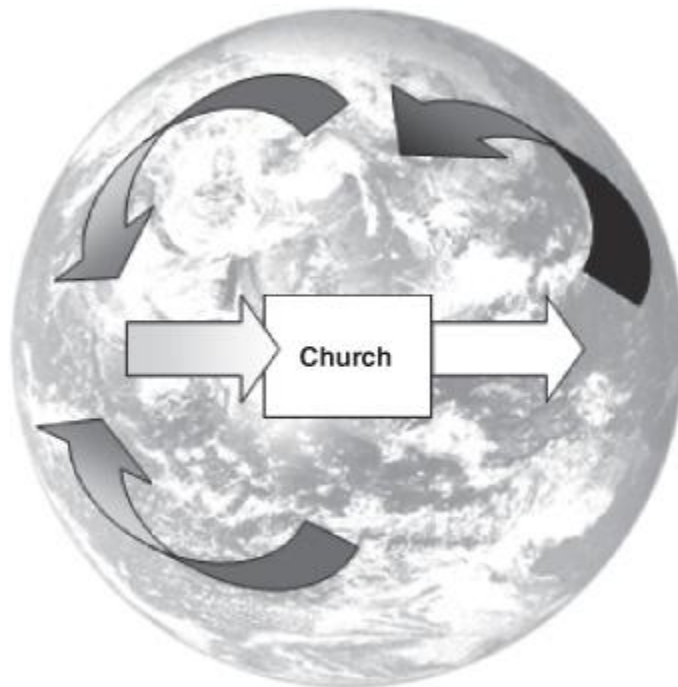


FIGURE 37-4 Church can cause world to become brighter according to Bishop Christopher.

Early on, we got the team to discuss the purpose of the UMC. Their stated purpose was “Make Disciples of Jesus Christ for the transformation of the World.” Lisa and I broke them into groups and had them draw a picture of their system.

Bishop Sharon Brown Christopher headed one group. Bishop Christopher and her group drew a picture of the Earth as viewed from outer space. The picture is provided in [Fig. 37-4](#). They represented the Church by a box and had dimly lit souls going into the box and bright souls coming out (some of these are my words, not hers). The flow into the box pulled other souls into the box and all were recycled back into the box (they did not stay lit but needed recharging). The group said if the Church were doing its job, then the world would get brighter and brighter.

Holy Mackerel. . . what a cool depiction of an overall system; it was very simple and easy to understand. This group had been exposed to a lot of TOC training so Lisa and I pressed them—“What is in the box and where is the physical constraint?”

Ezra Earl and his team had already done some work in that area and said that the four steps in the box (the Church) were:

1. Invite people into the box.
2. The people develop a relationship with God and with each other.
3. That relationship is nurtured by Bible study, prayer, etc.

4. People are sent out into the world to engage in God's work concerning injustice, mercy, and sharing the good news.

These four steps combine to make the process for forming a Disciple; just like specific manufacturing equipment combines to form a process that produces a car part. Although each of the four steps is an "operation" in the overall process, each step can be quite complex on its own.

"Okay, where is the constraint?" we asked. I thought that would be a hard question for them. The team did not even hesitate—"Step 2, developing a relationship with God." I was taken aback with the speed and certainty of their answer. They even mentioned that this is the case in 9 out of 10 churches. However, that inspired another question: "So how does that happen? How does one establish a relationship with God?" I think it is important to note that we have now made this extremely complex system a bit less complex by focusing in on one of the four overall operational steps. This group of Bishops is now immersed in describing Step 2—the people develop a relationship with God and with each other.

The group discussed and pondered that question for a while. In all of their discussions on how one establishes a relationship with God, this thing called a "spiritual leader" kept coming up. These leaders of the church agreed that the thing they called a spiritual leader was the key to helping create an environment in which an individual and God could better connect; therefore, "spiritual leadership" is the constraint of the UMC. We now need to look at "spiritual leadership" as we would any other skill set or piece of equipment, especially since it is our precious and most valuable resource (the constraint). We first needed to identify what it looks like.

"So if I walked through your factory (the Church) and tripped over a spiritual leader, what would he or she look like?" I asked.

I sensed that was an uncomfortable question. The participants came up with a description of a spiritual leader as one who is:

1. Humble and worships the Lord with joy.
2. Involved in daily prayer, Bible study, and devotionals.
3. Involved with others on a routine basis to discuss how God is working in their lives and to hold one another accountable.
4. Participating in acts of mercy and addressing injustice.
5. Telling others their faith story.

What they had described was a Disciple and had discovered that “Disciples make Disciples.” Another thing that they discovered is that they were paying a whole lot of folks to be spiritual leaders who did not meet that description. That particular issue was a problem but not one that the group was ready to address. There was enough work and progress to be made in focusing on the spiritual leaders they did have.

If spiritual leadership is the system’s constraint, then how do you “exploit” the system’s constraint? If you have 40 hours per week of operating time for a spiritual leader, what do you want them doing? Should they be attending meetings, washing windows, answering the phone and dealing with lawsuits? The answer to those questions is “no.” Those activities do not move the UMC closer to achieving its purpose. The Bishops defined the five steps in the description of a spiritual leader as “being on the path” and said that spiritual leaders should be on the path with one another and with their congregations executing and facilitating the four steps in the box (the Church).

Therefore, the major injection provided in the Evaporating Cloud of the core problem was, “Spiritual leaders are on the path with one another.” This is actually a very practical solution and doable. The Bishops could be “on the path” with their cabinets (staff), the staff could be on the path with those pastors who reported to them, and so on all the way down to the individual church member. There was excitement in the group. Each Bishop was in charge of a region of the United States, which they referred to as an “annual conference.” We began to formulate plans where each Bishop would head their annual conference in this plan going forward. It was going to be a very large challenge, as much of what a Bishop does has nothing to do with spiritual issues but with issues of running a large organization. They find themselves absorbed in legal matters and administrative minutia—just like too many CEOs of corporations. However, there was a lot of energy and perhaps they might be able to pull off a major change initiative. Again, I think it is important to pause and understand that we have taken a system as complex as the UMC with over 8 million members and boiled it down to a relatively simple process on what needs to be done to have a major impact on UMC’s ability to achieve its purpose.

Then the workshop participants’ “mindsets” became their major obstacle to progress. Ezra Earl called Lisa and me and said that he wanted to have breakfast with us because he was concerned that things were not going well. At our breakfast he scolded us and told us that we were putting too much emphasis on the Bishops and that we needed to treat everyone equally. Lisa and I argued that the Bishops were the ones in charge of their regions and should be the ones carrying the torch. He told us to back off and take us down a different path on the last days, which Lisa and I did. In spite of our best efforts, however, momentum died, and you could almost feel it prior to the end of our time together. Ezra Earl scheduled another breakfast with us and said, “I messed up,

didn't I?" He was not going to get any argument out of Lisa and me—yep, he did. We should have pressed ahead with the plan. There was a lot of learning and exchange of ideas; however, no implementation effort was forthcoming out of the session. I have always admired Ezra Earl for coming back and admitting his error. I have met very few people that had that kind of courage and leadership. It enhanced learning going forward.

Over a decade passed and Bishop John Schol, the newly appointed Bishop of the Baltimore-Washington Conference, the largest in the UMC, was attending a meeting of Bishops and the session that we conducted was being discussed. He learned that I was a member of his conference and after several meetings Chesapeake was hired to help them through a major transformation. What made it inviting to me is that Bishop Schol's plan looked like it had been copied directly from what we had discussed and developed 10 years earlier. I guess things are done in God's time, not ours. He called his initiative "The Discipleship Adventure," and it is almost exactly what we helped the Bishops develop years earlier. The elements of required leadership action are exactly the same five items as listed previously. However, Bishop Schol was ready to act. Hands-on implementation with this group lasted a year and they continued to implement after Chesapeake was gone.

One of the first problems the Bishop faced was being able to find time to be a spiritual leader (on the path) himself. When we reorganized the conference, we created a new position of Chief Operating Officer (COO). This individual would handle most of the legal matters and minutia of day-to-day operations, freeing the Bishop to focus more on making Disciples and being on the path with his leaders. We created "Discipler Groups" that were groups of pastors who met on a routine basis to practice the five steps together as leadership behaviors, as they led the way in executing the four steps in the box (the Church). While other UMC

conferences were reducing the number of people who “ministered to ministers,” the Baltimore-Washington Conference was increasing these numbers. These new leaders were called “Discipler Guides.” The role of the Discipler Guide was to facilitate the pastors in the conference being on “the path” with one another. There was some initial role conflict with the Guides and the traditional District Superintendent, but that was worked out over time. These changes took place in 2007.

How do you measure “Disciple making?” The Baltimore-Washington Conference came up with the following metrics:

1. Worship attendance.
2. Whether a church met their financial obligation to the conference.
3. The percentage of people who attended worship that were involved in small groups.
4. The percentage of people who attended worship that were involved in some sort of mission or service work.
5. The number of people who joined the church as a profession of faith.

The Methodist Church defined:

- The system to be analyzed as an “Annual Conference” or specific area of the country.
- The purpose of the system as to make Disciples.
- The measurement system as the five metrics listed previously.
- The system’s constraint as spiritual leadership.
- Exploitation of the system’s constraint as spiritual leaders being on the path with one another and their

flock while a COO manages church financial and legal affairs.

Results after Two Years

Is it working? This chapter was written in 2009. The Bishop has set as a goal the year 2010 when the negative trend stops and 2012 as the year all trends are on an upswing. In two of the four regions of the conference, they are already seeing positive results, so those two regions are ahead of schedule. If they stay the course, I am confident they will be successful.

Summary

Dealing with complex systems is fun when an organized systems approach is taken. Here are some of the things that have worked for me over the years.

1. One must first define the system. What are its boundaries? It has been my experience that the initial perspective of what you call “the system” will change.
2. Define the purpose of the system and how you measure success. If you can measure how to make Disciples, surely you can measure anything.
3. Remember that systems (and their cultures) are the combination of purpose (processes), relationships, and information flow.
4. Information flows through relationships so you can assume that if the relationships are improved so will information flow.

5. Begin globally and work from the outside in. What are the global processes that achieve purpose? What is the information required to achieve purpose?
6. Where is the physical constraint and is it in a desirable location? If not, take action to move it.
7. What obstacles exist that would prevent exploitation of the constraint?
8. Who needs to be involved to implement the change and what do they need to experience to change their mindset?
9. Never give up. Change is not linear and can accelerate at any time. Stay the course and be persistent.

There is no “cookbook” for addressing the problems in complex systems. If anyone says that there is, I would advise holding on to your wallet. There is no substitute for real people who have the knowledge, skill, and desire to address the complexity. Someone who has an understanding of the science of systems is going to be a necessity. In my book, *Enterprise Fitness*, I (Covington, 2009, 134) emphasize the importance of leadership in this role. This person needs to disrupt, honor, and align constantly during the change process. If the top leader in an organization is not ready to change in a complex system, leave and move on to the next system.

Reference

Covington, J. 2009. *Enterprise Fitness*. Mustang, OK: Tate Publishing & Enterprises.

About the Author

John Covington is president and founder of Chesapeake Consulting (CCI). CCI specializes in process improvement and leadership development in both commercial and government markets and has been in business since 1988. John did his undergraduate work at the U.S. Naval Academy and the University of Alabama, earning a BS in chemical engineering. Prior to starting CCI, John held engineering, management, and executive positions at a variety of companies including DuPont, Sherwin-Williams, Stauffer Chemicals, and several midsized paint companies. John is a Fellow of both the College of Engineering and the Department of Chemical and Biological Engineering at the University of Alabama. John is active in charity work for the developmentally disadvantaged and is an active member of his church. He enjoys biking, hiking, and training his German Shepherd, Maggie. He has been married to his wife Linda since 1972.

CHAPTER 38

Theory of Constraints for Personal Productivity/Dilemmas¹

James F. Cox III and John G. Schleier, Jr.

Introduction: A Status Report

Some people are very effective at their jobs and in their personal lives, while others never seem to be able to keep up in either. There are literally thousands of self-help books and articles discussing how to speed read; organize your home, your office, and your life; how to remember names and faces, numbers, etc. For almost every aspect of your life, there are books on how to improve. There is an awful lot of data and little information of value for the individual. We have this personal productivity chapter positioned in the complexity system section of this handbook for this reason. In keeping with the tenants of *Theory of Constraints* (TOC), we want to identify a few control points in managing your personal productivity that we hope will have significant impact on your ability to achieve your life goals and have a happy and fulfilling life.

The purposes of this chapter are to provide guidance in using the *Evaporating Cloud* (EC) technique to resolve chronic conflicts in managing both internal and external conflicts to achieve life's goals; understanding the differences between necessary conditions and goals; establishing personal life goals and supporting objectives;

understanding how to measure your progress toward these supporting objectives and ultimately the goals; knowing how to record and analyze how you use your time; understanding how to use priority planning, capacity planning, priority control, and capacity control to achieve your supporting objectives; and knowing how to use *Buffer Management* (BM) to improve your execution effectiveness. We also provide an in-depth application of using the *Thinking Processes* (TP) to achieve your life goal. We also feel that using the tools to plan and to control personal lives is fundamental to learning how to apply them in other environments.

Copyright © 2010 by James F. Cox, III and John G. Schleier, Jr.

Resolving Chronic Conflicts and Developing Win-Win Solutions

Dr. Eliyahu M. Goldratt has developed the EC thinking technique to assist in identifying and solving both day-to-day and chronic conflicts in businesses (Goldratt, 1993; 1994; 1995). At a Jonah Upgrade Workshop, Effrat Goldratt and Lamor Winter (1996) describe the use of the three-cloud approach to building a *Current Reality Tree* (CRT) as it applies to individuals. They had invented this approach, testing it out in workshops, and later proposing that it be used in organizations. Resolving personal external and internal conflicts/dilemmas is a major factor in improving personal productivity.² You cannot focus your efforts if you do not know what the problem is that is blocking you from achieving your goal(s). While most books and chapters on personal productivity ignore conflict problems as topics, we feel that acquiring skills to solve these problems provides the foundation for both personal and white-collar

productivity, managing many thought activities, and overall managing the organizational improvement processes. Personal and white-collar productivity requires focus, concentration, and motivation. The elimination of conflict problems that block or inhibit applying these factors to a problem is a necessary condition to being productive.

Most improvement books discuss “to-do” lists, but few discuss linking daily activities to short-term objectives to life goals. Moreover, fewer discuss developing detailed plans to change your life. The *Negative Branch Reservation* (NBR) is useful in testing the impact of actions in a good solution. The *Prerequisite Tree* (PRT) is useful in identifying and overcoming obstacles in implementing your solutions. Each of Dr. Goldratt’s techniques provides a graphic display of the logical relationships surrounding a problem. In Chapters 24 and 25 of this handbook, Oded Cohen and Lisa Scheinkopf describe the procedures for both constructing and communicating these applications. We will not duplicate their efforts. In this chapter on personal productivity, we will present a couple of applications of these two simple but highly effective TP to demonstrate their application to personal problems. Both tools provide the basis for understanding other TP presented in other chapters and in the detailed application in this chapter. You will be quite surprised at how these TP allow you to verbalize your intuition and how useful they are in identifying and communicating ideas to other people. The personal productivity application in this chapter shows how the full TP assisted one student to achieve a life-long dream.

You face numerous personal dilemmas over the course of your life. These dilemmas sap your energy, concentration, focus, motivation, etc. far more than you realize. Some might be one-time situations where the decision can change the whole course of your life; some might be a series of recurring dilemmas that evolve into a chronic conflict between you and another party; and some might be just

plain and simple, nagging, day-to-day dilemmas. Let us provide a series of dilemmas drawn from personal experiences. Imagine the impact these dilemmas have on your energy, concentration, focus, and motivation in all facets of your life. We use one of our sons as the example. Please recognize that a similar series of dilemmas can exist with your daughter, spouse, parents, siblings, coworkers, subordinates, supervisors, etc. The objective (the A) and requirements (the B and C) may occasionally change for the person and the relationships you have with them, but this segment of the cloud usually repeats itself again and again across a number of different dilemmas.

Background: Father-Son Dilemmas

My son and I (and prior to that my daughter and I) seemed to continually argue about nothing (as far as I was concerned). I kept saying, “No, are you crazy?” or “No, you are too young!” to most of his requests. I was frustrated with our degenerating relationship. I hated to always be the bad guy. I wanted him to be safe, honest, well mannered, hard-working, and motivated. I felt it was my responsibility to ensure that he grew up to be a model citizen.

I finally realized I was in a chronic conflict.³ At this point, I reminded myself to question my own and my child’s assumptions based on the current situation and try to come up with a win-win solution. When he entered high school, I found the need to diagram some of the conflicts to gain a better understanding of our declining relationship. Our common objective was simple: to maintain a lasting father-son relationship. I consider this father-son cloud to be a chronic conflict as the son will always be pushing for more and more freedom as he gets older and the father will always examine the situation for security and want the son to make the right decisions. If I continually say “No!” then I

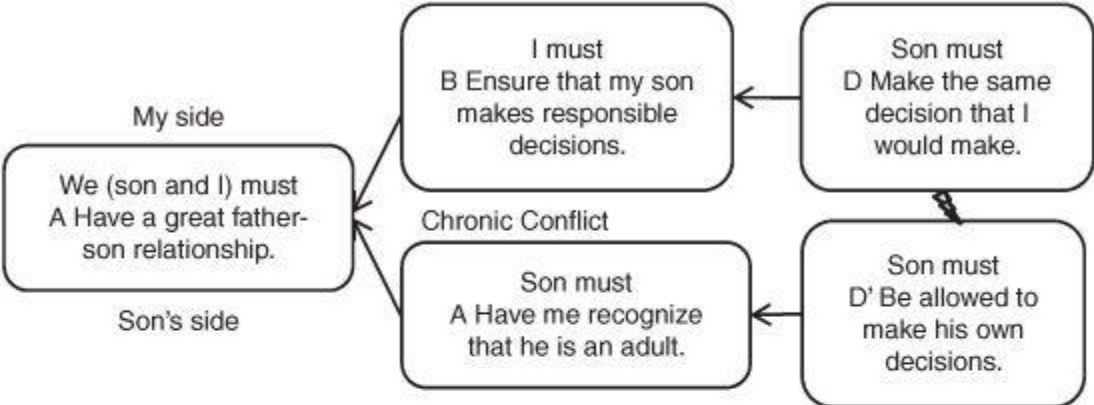
create a relationship that I will regret the rest of my life. My son will just do what he wants behind my back and our lines of communication will be damaged or broken. On the other hand, if I continually say “Yes,” then I am irresponsible and am neglecting my role as a parent to provide a safe and secure environment.

In order to have a great father-son relationship, I want to ensure that my son makes responsible decisions. In order to have a great father-son relationship, my son wants me to recognize he is an adult (recognize that even a 12-year-old thinks he is an adult). In order to have a great father-son relationship, we need to accomplish both—I want to ensure my son makes responsible decisions *and* my son wants me to recognize he is an adult. In order that I want my son to make responsible decisions, my son must make the same decisions I would make. However, in order that my son wants me to recognize he is an adult, my son must be allowed to make his own decisions. On one hand, my son must make the same decisions I would make, while on the other hand, my son must be allowed to make his own decisions. Are these two actions D and D’ in conflict? Yes. Please read the cloud in [Fig. 38-1a](#) carefully to understand each person’s assumptions and their suggested injection for this chronic conflict.

I will illustrate the chronic nature of this father-son relationship with several specific examples drawn from my personal relationship with my son. I am not saying that I have a perfect relationship with my son through this case study. I am saying that there is a chronic problem in the father-child (any parent—any child) relationship *and* it must be recognized and addressed as such *or* there is no relationship.

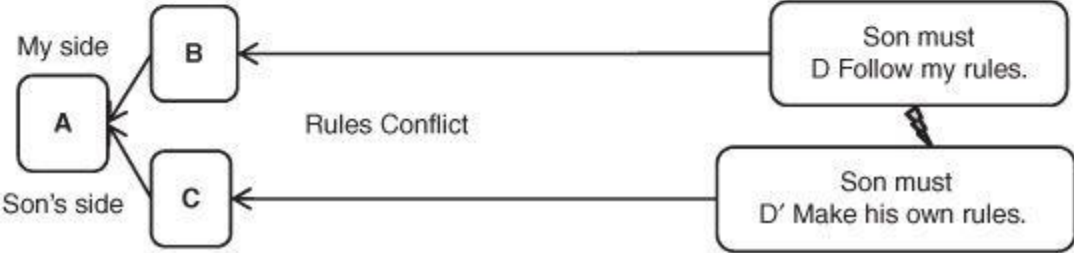
Father-Son “Rules” Dilemma (Primary School)

Situation: “Put your toys away.” “Clean your room.” “Make your bed.” “Clean up the bathroom.” “Go to your room and study.” Are these commands that you seem to be giving to your children more frequently? Are their actions coming slower and slower? My son was starting to question—Why? Why do I have to do it now?



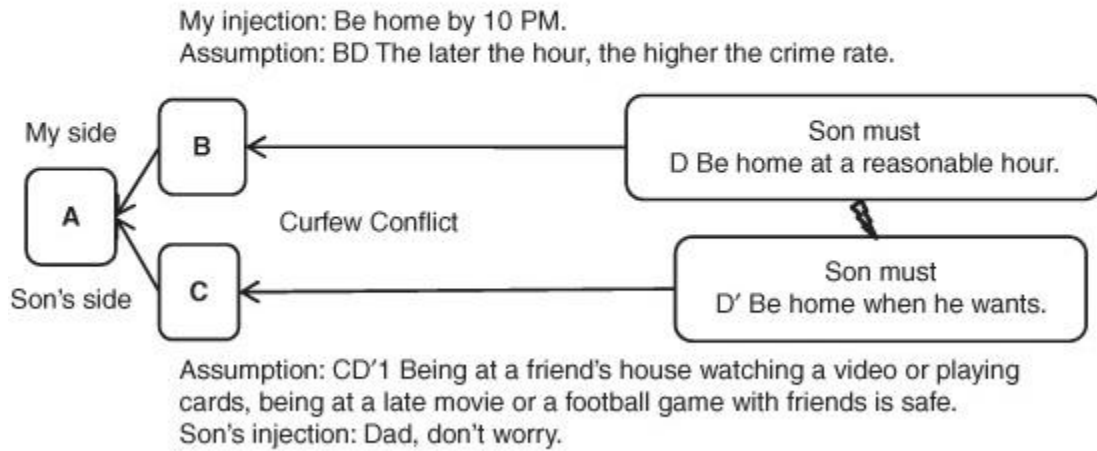
a. Father-son chronic conflict.

My injection: He can make his own rules as soon as he is 21 and paying his own way. Do what I say!
 Assumption: BD I am responsible for his behavior while he is a minor or while I am paying his way.



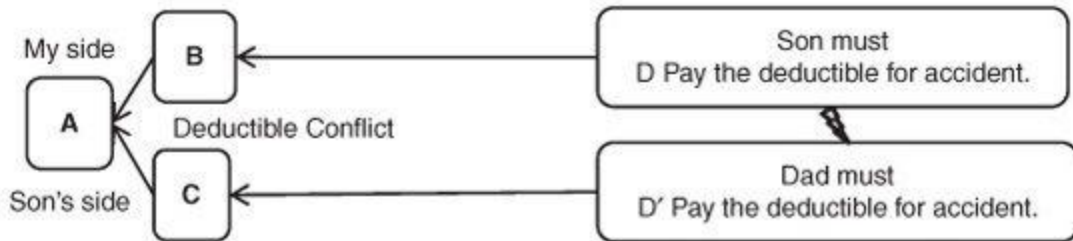
Assumption: CD'1 Son is responsible for his own behavior.
Son's injection: Ignore dad.

b. Father-son rules conflict.



c. Father-son curfew conflict.

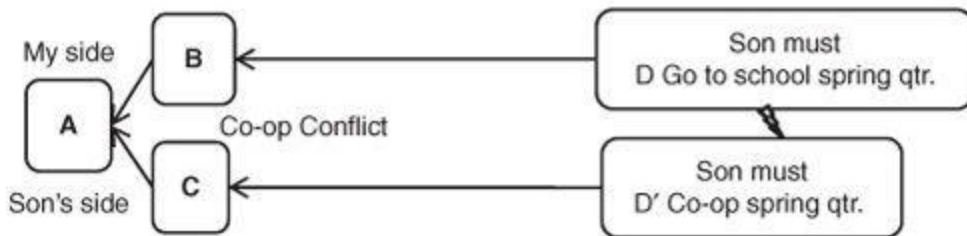
Injection: Take the money from your savings (CDs).
 Assumption: BD Son caused the accident; therefore, he pays.



Assumption: CD'1 Yes, I should pay. BUT you wanted me to tie up the summer with a course and trip instead of getting a job.
 Injection: Dad should pay the deductible for the accident.

d. Father-son deductible conflict.

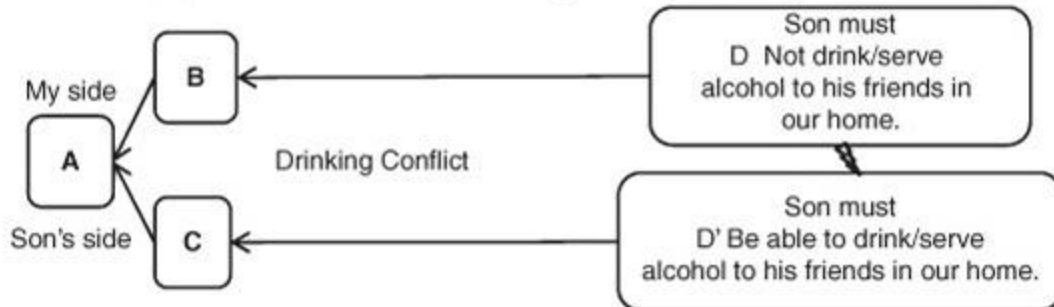
Injection: Check to see the impact of co-oping on graduation date.
 Assumption: BD Many course prerequisites exist in engineering.



Assumption: CD'1 My advisor told me it would take 2 years.
 Injection: Please recognize that I am an adult. Let me decide.

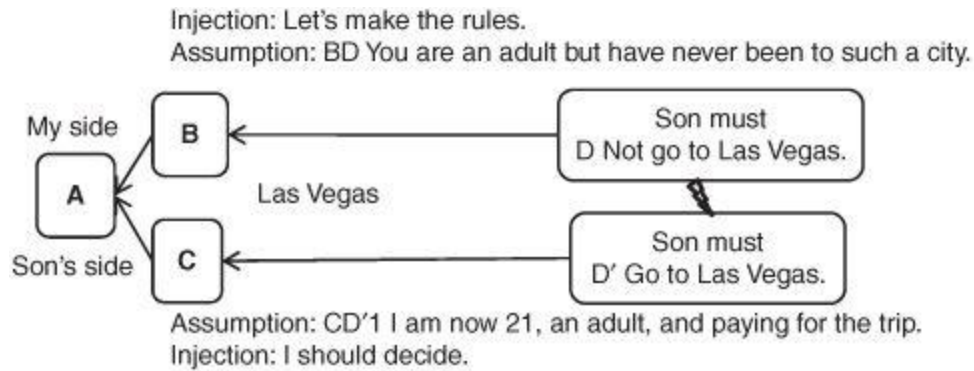
e. Father-son co-op conflict.

Injection: He is now 21 years old; some friends are 21; others are younger.
 Assumption: BD He is a minor; he can get in trouble serving alcohol to other minors.

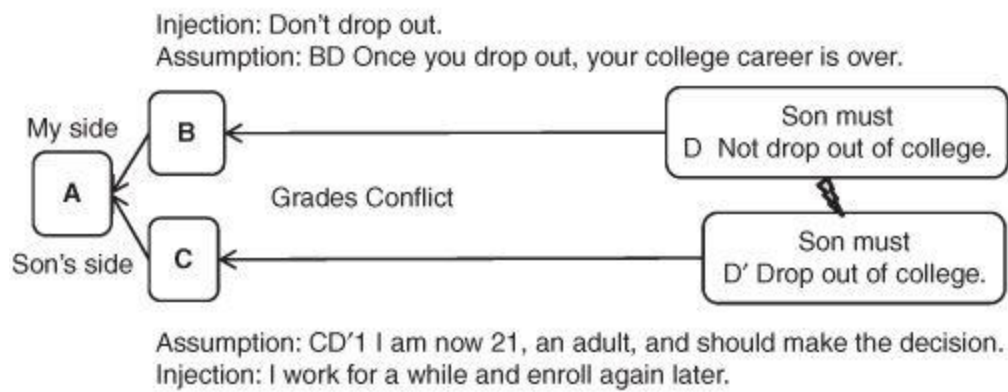


Assumption: CD'1 I am now 21, an adult, and should make the decision.
 Injection: My friends and I can go out to a bar and celebrate my 21st birthday.

f. Father-son drinking conflict.



g. Father-son Las Vegas conflict.



h. Father-son grades conflict.

FIGURE 38-1 Father-son relationship conflicts.

This questioning of rules marks the start of the chronic conflict of the father-son relationship. Be careful in how you respond as this period marks the beginning of the struggle between a parent and the child; of letting the child grow to accept responsibilities. Please read the cloud in [Fig. 38-1b](#) carefully to understand each person's assumptions and his suggested injection.

Win-win solution (based on the NBR): Dad has four basic rules (short version) which can never be violated—(1) No drugs. (2) No sex. (3) No smoking. (4) No drinking and driving. All other rules are negotiable based on the situation. I recognize that my son may have violated some of these rules occasionally, but I added a fifth rule when he reached the age of 21. The added rule is: (5) Use the cloud and NBR

for evaluating the decisions you make. Recognize that you must live with the negative consequences of the decisions you make.

Father-Son "Curfew" Dilemma (High School)

Situation: My son and I seemed to argue continually about what time he was required to be home. This situation worsened when he got his driver's license. Please read the cloud in [Fig. 38-1c](#) carefully to understand each person's assumptions and his suggested injection.

Win-win solution: Based on the situation, determine an approximate reasonable hour. If you change plans or may be late, then immediately call home and renegotiate. For any serious problem, call home immediately.

Father-Son "Major Issue" Dilemma (College)

Occasionally, a major issue will crop up between my son and me. When it does, it lasts a long time. The following is just such an issue.

Situation: My son was driving his truck and had an accident, which was his fault. He admitted that he caused the accident (fortunately, no one was hurt). The dilemma—our insurance had a \$500 deductible. I wanted him to pay the \$500 deductible and he wanted me to pay the \$500 deductible. Please read the cloud in [Fig. 38-1d](#) carefully to understand each person's assumptions and his suggested injection.

Win-win solution: My son pays the deductible by working for his dad (flexible hours). He cleaned and stained the deck. He also set up dad's computer system, checked all files and diskettes for viruses, and restructured the hard drive. He did several tasks dad never could find time for or the desire to do. Son desperately wanted to resolve this

situation but did not want to cash in a certificate of deposit (CD). He always saved his money to buy CDs and had never cashed any in.

Father-Son "Co-Op" Dilemma (College)

Situation: My son calls home very excited about the opportunity of co-opping with a company at \$15 per hour. We were quite excited for him as well. The job would look good on his resume and may result in a future job with that company. He had already checked with his advisor to determine the impact of co-opping on his graduation date. The advisor had stated that it would delay his graduation 2 years. He is in chemical engineering and several courses are specialized and offered only once a year. Understanding the problems of taking prerequisites, I was quite concerned. Please read the cloud in [Fig. 38-1e](#) carefully to understand each person's assumptions and his suggested injection.

Win-win solution: My son checked again with his advisor to determine exactly which courses remained to be taken, when they were offered, and when he could take them based on having the prerequisite courses for starting the co-op program in the spring, summer, and fall quarters. He found that it would add two years (go to school in the spring and fall and [start] co-op in the summer and winter quarters) to his current graduation date under the best situation and 3+ years under the worst case (co-op in the spring quarter, when the job was extended, go to summer school, and co-op in the fall). Given this information, he talked to the company personnel and decided the proposed co-op position was not a good situation at present.

Father-Son "Drinking Age" Dilemma (College)

Situation: My wife and I do not serve alcohol in our home. I occasionally drink socially. We do not serve or drink alcohol to set a good example for our children. How can I tell them not to drink and drive and then do it, or even worse, how can I serve alcohol to guests in my home knowing that they must drive home afterward? We have always told them, “Never drink and drive. And never let a friend drink and drive. Call us, we will pick you up at any hour, no questions asked.”

My son was to turn 21 on the weekend of the Auburn University (he went to AU) versus University of Georgia (I taught at UGA) football game. The game was in Athens, GA (our home) so he wanted to bring nine fraternity brothers to spend the weekend with him. He wanted to serve alcohol to them at his party in our home. We knew that they would party that weekend with or without our permission at local bars. Please read the cloud in [Fig. 38-1f](#) carefully to understand each person’s assumptions and his suggested injection.

Win-win solution: We developed a set of strict rules for serving alcohol in our home.

1. My son must enforce the rules—he is an adult and is responsible for the safety of his guests.
2. He will card each friend to ensure that he is 21. He will not let anyone below 21 drink.
3. He will monitor the drinking of each friend.
4. He will not let anyone who has been drinking drive.
5. He will inform his friends of these rules before they come into our home.

Father-Son “Las Vegas” Dilemma (College)

Situation: My son called home to see if he could go to Las Vegas with his friends. After listening patiently to the situation, we heard, “Can I go?” His roommates (nice young guys) were going to Las Vegas with one of the roommate’s father and uncle. They had rooms at a great hotel and it would be four college students and two adult males. Please read the cloud in [Fig. 38-1g](#) carefully to understand each person’s assumptions and his suggested injection.

Win-win solution. My son goes to Las Vegas with four rules.

1. He is going with three college friends and two parents.
2. He is staying at a great hotel.
3. He is paying for it.
4. He calls prior to leaving, at arrival, departure, and return to Auburn. He also calls immediately if any problems arise.

Father-Son “Poor Grades” Dilemma (College)

Situation: My son is a rising senior majoring in Chemical Engineering. He joined a fraternity his freshman year and at times I think he is more interested in fraternizing than studying. He started well academically and has continually declined in grades over the past three years. We have continually argued about his lack of studying. He continually blames his teachers for his poor grades. Being a teacher myself, I realize that some teachers are bad, but my son cannot be getting all of them. I blame him for his poor grades. I recognize that if he does not change his behavior he will not graduate, he will graduate but not be able to get

a job in his field, or he will get a job but be unable to function. Every situation looked bad to me. His future is not happy, how can I change the situation? Please read the cloud in [Fig. 38-1h](#) carefully to understand each person's assumptions and his suggested injection.

Win-win solution: By mutual agreement among AU, my son, and me, my son dropped out from college for six months and took two full-time jobs, one landscaping and mowing lawns all day and one as a bartender in the evenings. He figured out somehow between these two jobs that he was smarter than the brains required for either job. He went back to college, changed majors, worked part-time, improved his grades significantly, and graduated.

Father-Son Chronic Dilemma Summary

These reflections on the ECs of a father-son relationship are not meant to describe stellar decision-making. They are provided to illustrate the chronic nature of relationships among people whether the other person is a parent, child, sibling, friend, business associate, subordinate, peer, or supervisor. The ECs reflect what Covington (2009) describes as the elements that define relationships and culture in an organization:

1. Trust and honor people.
2. Our purpose and processes to achieve our purpose are clear.
3. Ongoing education and information.⁴

These same elements form the basis of sound, open, healthy relationships with other people.

The sooner we recognize that we are in a chronic conflict situation and identifies the common objective, each party's requirement, the opposing actions (usually this is clearly

defined), and the underlying assumptions for each side, the sooner one can study how to build a win-win lasting relationship with the other person.

Personal Productivity Dilemma—Where to Spend Your Time?

We examined the father-son chronic conflict and several specific examples to illustrate the complexity of only one dimension of one facet of a person's life. Let's look at other chronic conflicts that you have faced or are currently facing in hopes of improving your productivity. Prior to doing that, let's first examine the EC template and some helpful hints on constructing the EC.

A Review of Constructing the Evaporating Clouds

A number of other chapters provide detailed instructions on how to construct the EC with examples of each step. This review is not meant as a comprehensive procedure of how to construct and test the EC logic. Study [Fig. 38-2](#) to ensure that you remember and understand how to construct ECs. Answer each of the questions in the blocks in the template, using complete sentences in the order suggested. Follow the guidelines and helpful hints in the figure. Check the logic by reading the cloud using necessary condition logic: In order for ___ to (A, B; A, C)____, then (B, D: C, D')____. Do the statements make sense? Surface and build the assumptions. We will assume you know how to construct valid ECs in the remainder of this chapter. If you feel you need additional instruction, then go to Chapters 24 and 25 on the TP.

College Student Dilemma (Undergraduate)

In addition to having chronic conflicts with people you interact with, you can also have chronic conflicts with yourself. The next example is one such situation. We call it the classic dilemma of a college student.

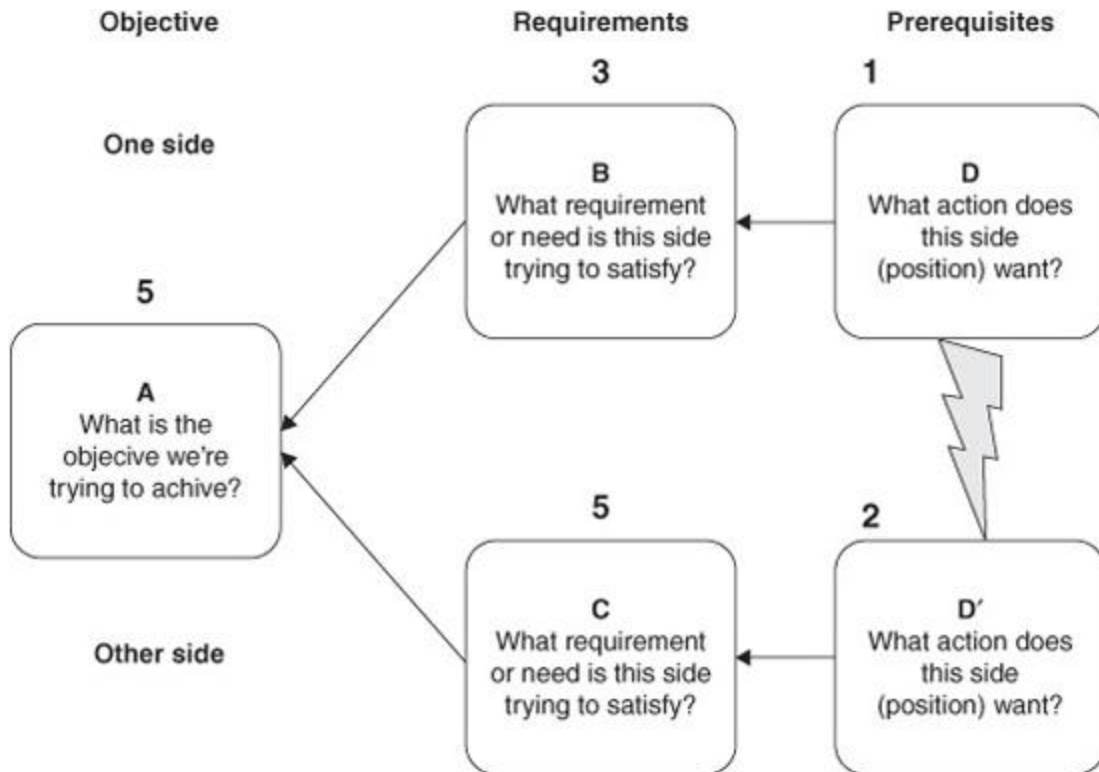
Situation: Most college students go off to school and for the first time must plan, execute, and control their own daily lives. Some are successful; many are challenged; some fail miserably. Students have many demands for their time: classes, labs, studying, eating, sleeping, exercise, part-time or full-time jobs, and leisure activities such as playing sports with friends, attending college sporting events, college plays, movies, partying, playing cards, watching TV, etc. (the list is almost limitless). With so many choices, how does a student make the right decisions in allocating and using their time? The dilemma breaks down into doing the things I must do to succeed in college versus doing the things that I want to do to enjoy the college experience. Read [Fig. 38-3](#) carefully.

Notice in the situation description, we provide the functions of planning, executing, and controlling. Another key function is prioritizing since, while in college, as in life, we never have enough time for all we have to do and want to do. Let's examine this life situation in more detail before getting into the direction of the solution.

EC of the Classic Dilemma of White-Collar Burnout

After college life and its dilemmas, careers take center stage. Many college graduates start their new business career with a bang by putting in 60 to 70 hours per week. Their initial and continued efforts are rewarded with raises and promotions but the stress, frustrations, and continued high pressure to perform force them to sacrifice their personal, family, social, and professional lives to gain security in an unsecure environment. Finally, one day they

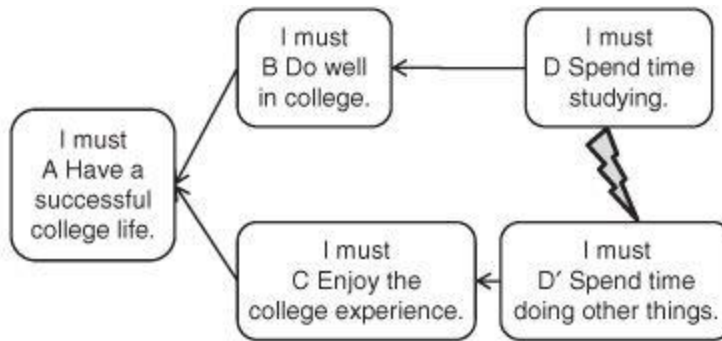
look back over their lives and in retrospect wonder where their lives have gone astray. They are worn out, insecure, and ready to get off the treadmill. This situation is known as “white-collar burnout.” The situation is depicted in [Fig. 38-4](#) as an EC along with our underlying assumptions for the causal relationships. The objective [A] of the young graduate is to have a satisfying life. The graduate feels that he must [B] achieve his life goals and simultaneously [C] meet the necessary conditions of his life. Both requirements (B and C) require that the graduate devotes time, motivation, concentration, effort, and energy to achieve the requirements. Of course, the dilemma then is that there is not enough time to devote time to everything. The underlying assumptions provide the logic. The burnout usually comes when the graduate recognizes the amount of time, energy, etc. devoted to work to the neglect of the other areas in his life. Let’s examine each part of the cloud and this dilemma in more detail, but first let us review some basic TOC concepts and apply them to personal productivity. Once you have read the remainder of this chapter, you should revisit these clouds and assumptions to determine how you are poised to achieve a happy and satisfying life.



Helpful hints for constructing a valid cloud

- Answer the five questions provided in the EC template based on the storyline.
- Answer the questions from right to left—D first, D' second, B third, C fourth, and A fifth.
- Use simple but complete sentences as entities.
- Don't use compound verbs or more than one direct object in A, B, or C, as an assumption may then be true for one part of the entity and false for the other.
- Specify who must achieve the objective, each requirement, and each action.
- Check the logic of the conflict.
- Is the objective correct?
- Are both requirements needed to achieve the objective A?
- Can both B and C exist simultaneously?
- Is B (C) the requirement driving action D (D')?
- Both D and D' should be described as actions or proposed actions.
- If one takes action D, then this action jeopardizes requirement C.
- If one takes action D', then this action jeopardizes requirement B.
- Is the action described in D the opposite of the action described in D'?
- Use present tense in all entities and assumptions.
- Assumptions should not contain conditional logic (e.g., if or because)
- Is the assumption valid (true)?
- Does it exist in the current reality (present environment)?

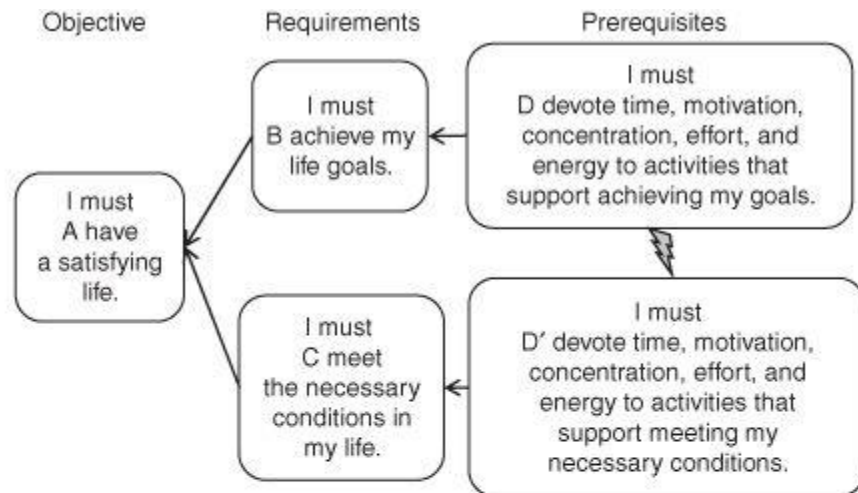
FIGURE 38-2 EC template and helpful hints.



Assumptions:

- AB Doing well in college provides the foundation for doing well in my career.
- AC1 The college experience includes making new friends and developing my social skills.
- AC2 The college experience is the last hurrah before full-time life.
- BD1 There are no shortcuts to learning.
- BD2 College is very difficult.
- BD3 I have always had to study hard to get good grades.
- CD'1 All work and no play makes Jack a dull boy.
- CD'2 I must work to put myself through school; work is a necessary condition requiring time.
- CD'3 I must maintain my sanity.
- DD'1 I can't do both satisfactorily as there are only 24 hours in a day.
- DD'2 I can't do both simultaneously as they are independent of each other.

FIGURE 38-3 EC and assumptions of the classic dilemma of a college student.



Assumptions:

- AB1 Everyone expects me to do well.
- AB2 Achieving my goals (no matter what they are) brings satisfaction.
- AC Without performing the necessary conditions of my life, the life is disrupted.
- BD1 Time, motivation, concentration, effort, and energy are all required to achieve life goals efficiently and effectively.
- BD2 No shortcuts exist in achieving life goals.
- CD'1 Necessary conditions are just necessary to sustain my life as I define it.
- CD'2 Any missing factor (time, motivation, concentration, effort, and energy) translates into spending more time on the activity.
- DD'1 I can't do both satisfactorily as there are only 24 hours in a day.
- DD'2 I can't do both simultaneously as they are independent of each other.

FIGURE 38-4 EC and assumptions of the white-collar burnout dilemma.

Personal Productivity—Establishing Goals, Strategies, Objectives, Action Plans, and Performance Measures

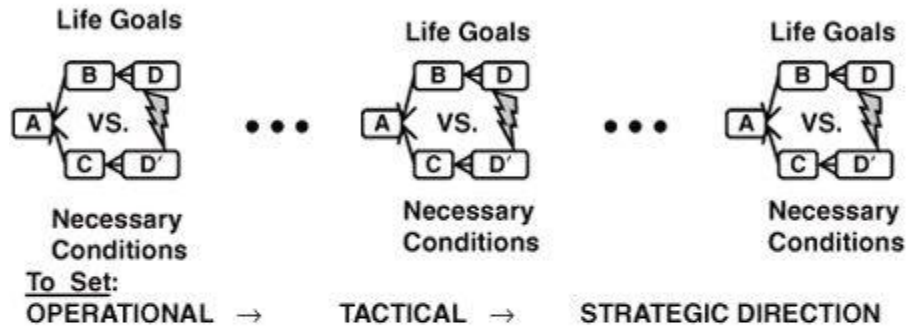
In [Fig. 38-5](#), we present an overview of the facets of one's life, and how they relate to each other to improve our personal productivity. Our definition of *personal productivity* is *moving towards achieving our life goals*. Most individuals are in a firefighting mode, moving from one crisis to another in each facet of their life. To move out of the firefighting mode, you have to identify and use tools that allow you to focus on one or two tasks at a time that move you toward

achieving your life goals in that facet. In some instances, you are in a chaotic environment where constant firefighting is the norm. In the TOC vernacular, this is the “What to Change” environment and the direction of “what to change to” provided by the TP helps the individual to determine what is important in your life. You must first find and use tools to move you to a stable environment. You must spend quiet time examining the five facets of your life: personal, family, friends/community, work, and professional. For each facet, a number of dimensions might exist; for example, in the personal facet, you might have goals (or necessary conditions) for physical, mental, and spiritual dimensions. You must decide what is important in each facet and dimension in the short and long term. However, a goal is only a dream unless you develop a plan and schedule for achieving it, then execute the schedule, and control interruptions to ensure task completion. The plan must provide the strategic direction for achieving the goal and supporting objectives and measurements that indicate your progress toward your goal. The plan must link from the strategic direction to the shorter term (tactical objectives) to the day-today activities that make up the operational plan. This operational plan (the “to-do” list) provides the mechanism: “How to Cause the Change?” Each day a “to-do” list of actions should include tasks that move you toward your supporting objectives in each facet of your life. These actions should provide progress toward weekly, monthly, yearly supporting objectives. Progress toward each supporting objective should be measurable and measured frequently to provide feedback. This feedback should be used to determine whether the supporting objective was achieved, whether the action was useful in moving toward the supporting objective, whether different actions are now required, etc. Each part of this diagram is discussed in detail next.

GETTING THERE!

From where we are now → To → Where we want to be!
 Firefighting To Focus on Life Goals

What to Change? → → → What to Change to?
 (To Go From: CHAOS ... → STABILITY... → GROWTH... → ACHIEVEMENT)



How to Cause the Change?

<u>Now Actions</u>	<u>Supporting Objectives</u>	<u>Supporting Objectives</u>	<u>Supporting Objectives</u>	<u>Life Goals</u>
TO DO List Actions	TO DO List Actions	TO DO List Actions	TO DO List Actions	TO DO List Actions
Personal ...	Personal	Personal	Personal	Personal
Family ...	Family	Family	Family	Family
Friends ...	Friends	Friends	Friends	Friends
Work ...	Work	Work	Work	Work
Professional ...	Professional	Professional	Professional	Professional
<u>With Measures:</u> Daily...	Weekly...	Monthly...	Yearly....	

FIGURE 38-5 What to Change, What to Change to, and How to Cause the Change in personal productivity.

What are your goals in life? Most of you want to succeed in business or you would not be reading this handbook. Some of you may change your mind about a business career after reading this chapter. Some of you may change your mind after a few years in business. You will probably have 50 years in the job market! You have objectively determined you wanted a business career, or you just wandered into the

business school not knowing where you wanted to go, or you graduated in another curriculum and ended up working in business. Maybe your Dad, Mom, or a brother or sister influenced you to select a business career. It may be the right decision or it may be the wrong decision for you.

Goal setting demands considerable time and concentration. You need to reflect on what you like to do. Do you like to interact with people? Do you like the sense of accomplishment derived from helping someone? Do you like to work with young kids? Do you like to solve computer problems? Do you want to go into the family business? Do you have a few or many friends? How involved are you in community activities? What do you want to do with your life? What are your goals?

Goal setting should take place in five different facets of your life—personal, family, friends/community, work, and professional. You need some direction, some goal, so you have an idea of where you are going in each and all dimensions of your life to be able to balance your time across these facets to achieve your goals. Most of you are responding to events in each of these facets each day. However, you should be seeking activities that move you toward your goal in each facet. What goals are you trying to accomplish in each facet? Dimensions of your personal goals include physical, mental, and spiritual. Dimensions of your family goals include your relationships with Mom, Dad, siblings, spouse, kids, and community. Dimensions of your work goals include current projects, pay, and work environment. Dimensions of your professional goals include higher degrees, certifications, and new skills development. Do not forget that you will be on the job market for 50 years and your current skills may be obsolete in a few years.

What is a goal versus a necessary condition in your life? A goal is generally viewed as something where more of the goal units are better. Making more money now and in the future means continually striving for improvement. A

necessary condition means that some amount of an item is satisfactory to you, more is not necessary. A goal may be to get an A in this course, while a necessary condition may be that you must get at least a C in order to take the advanced courses. There is a big difference in the amount of time, effort, concentration, and motivation required to achieve an A versus a C.

A goal for one person may be a necessary condition for another. For one person, a goal may be to run less than 8-minute miles in a marathon, while for another person, a necessary condition is to walk briskly for an hour three times a week. Doing well in the marathon means a lower time is better. You may consider walking three times a week as the minimum amount (the necessary condition) for maintaining your physical fitness.

For one person, a work goal may be to find a job where you can make as much money as you can. On the other hand, for another person, a necessary condition might be to find a job where you make at least \$40,000 annually, but work in the outdoors. More money may not be important to you—\$40,000 is enough for you to live the lifestyle you want. One person's goal might be another person's necessary condition. You must recognize in each facet of your life what is a goal and what is a necessary condition. A goal may change into a necessary condition in the short-term and then change back. Suppose you have set a goal of losing 20 pounds over the next six months. You have lost 12 pounds thus far, but Christmas is approaching and you want to enjoy the holidays with family. You may decide to "maintain my weight" until after New Year's instead of forcing yourself to diet over the holidays. After New Year's you are back to the diet and trying to hit your target of 20 pounds. Knowing the differences between goals and necessary conditions (reduce frustration) are important, so that you know where to expend your focus, concentration, motivation, effort, and time.

In a work or school environment, recognizing the differences between goals and necessary conditions of people, and the differences in the actions of those people, is vital to understanding teamwork and reducing your frustration level. For example, you have probably worked on class projects as a team of three or four students. Sometimes you have a teammate who really works hard and sometimes you have a teammate who does not seem to care. The difference may be that one teammate views the project as the means of achieving the goal of an A in the course, while another views the project as a means of achieving the necessary condition of a C in the course. Their level of activity (time commitment, motivation, concentration, effort, and energy) supports their objective for the course. A good question to ask potential team members prior to forming the team is, "What grade are you going to work for on this project?"

One last point concerning goal setting is the understanding that goals in each facet of life can, and do, change. If you graduate from a school in business and after a couple years find that you dislike business and want to do something else, spend some time evaluating where you are and where you want to go. Have your interests changed? You have your whole life in front of you. Many workers today dislike their job or their work environment, but fail to recognize their ability to change. You should enjoy each facet of your life, and if you encounter obstacles, address them. Reassessing your goals and developing new goals in each facet of life is an important part of your continuous improvement process.

To manage your time effectively, you have to know where you are headed. That is, you should establish both long-term goals and supporting short-term objectives. You must be proactive in setting and achieving your goals. They provide you a direction for focusing your daily personal and professional efforts. Strategy tells you *how* you are going to

accomplish your goals. In addition to these goals, you should set daily, weekly, monthly, and quarterly objectives that move you toward achieving your long-term goals. You must develop a strategy and supporting action plans to achieve your goals and objectives. Your goals and objectives should be defined so that you can measure your progress toward achieving them. Measuring progress requires designing a performance measurement system that consists of performance criteria, performance standards, and performance measures. A *performance criterion* is a factor to be evaluated, a *performance standard* is the desired or acceptable level of performance, and a *performance measure* is the actual performance. The steps in establishing goals, objectives, and a measurement system are provided next.

1. Identify your long-term goal and its supporting shorter-term objectives.
2. Develop a strategy (how you are going to accomplish the goal and objectives) and supporting action plans to then accomplish these things.
3. Identify a performance criterion for evaluating your progress toward your short-term goals and objectives. (What am I going to measure?)
4. Identify short-term standards for your performance criterion that reflect meeting your objectives and a longer-term standard that reflects goal attainment.
5. Monitor your progress by measuring performance on your short-term objectives.
6. Compare your performance measures to your performance standards.

7. Take corrective action, if necessary.

Let's apply these steps to a specific situation where the student is a little more mature in approaching schoolwork than the typical undergraduate student described in the EC in [Fig. 38-3](#). Suppose you work full time and are a full-time MBA student in a night program.⁵ Your professional goal is to [A] Graduate with honors from an MBA program while not sacrificing the other facets of your life. For your strategy (how to accomplish the goal), you decide to re-evaluate the time commitments made to various life facets where you can devote more time to MBA studies so that graduating with merit can be achieved. Examining the generic cloud in [Fig. 38-3](#), the objective [A] remains the same, your [B] and [C] requirements are: [B] Achieve honors in the MBA program (an A-average across all courses) and [C] Satisfy the other facets of your life while completing the MBA program. This requirement encompasses work, family, friends, etc. Of course, the D—D' dilemma is the same: You do not have enough time to do both simultaneously.

You have decided that the performance criterion for your MBA degree is the course grade. As you take exams and turn in projects in each course, you are able to chart your actual progress against your short-term standard for the course grade. Charting your performance measure (actual grades) against your performance standard (desired course grades) indicates your progress toward your long-term goal of graduating with merit. Are you on schedule, ahead of schedule, or behind schedule with your average grade? You might want to re-evaluate your course objectives for next term based on your results from this term.

When a deviation (difference between standard and actual grades) in grades occurs, identify the cause. How can you address the cause? Suppose that you were on a project

team where the others did not contribute their share to the project. How might you address this situation in the future?

What to Change—How Do You Currently Use Your Time?

Most people recognize that they do not use their time effectively in accomplishing their goals. Many read self-improvement books on how to organize, how to improve memory, how to speed-read, etc. They try to improve everything instead of focusing on the core problem. Is time the problem with you?

It takes time to accomplish your short-term objectives and goals. Do you know how you *use* your time? Do you spend it freely or manage it? Are you reactive or proactive in managing your time? Do you plan your time or just let things happen? Once your time is planned, do you immediately abandon your plan once a disruption occurs? Do you ever accomplish your goals and objectives?

Prior to deciding on a plan for personal improvement, you should find out how you currently spend your time. What are the UDEs of poor time management? Set up a time analysis form and run enough copies to cover a week of recording your activities. A time analysis form is quite easy to construct in a spreadsheet. At the top of this form, write your objectives for the day (a simple “to-do” list). Some of these daily activities should support your short-term objectives. If you do not currently plan your day, leave the objectives section blank. Most people do not use “to-do” lists. In the first column, enter the time you usually get up (6:00; 6:30; 7:00; etc.) and 30-minute intervals to when you usually go to bed. In the second column, enter the activity that you are performing during that 30-minute interval. Attempt to record how you spend your time every half hour or less. If you wait to record your activities until the end of

the day, you lose your perspective of the amount of time taken by various activities and you fail to record major interruptions. The third column, the comment column, allows you to provide additional insight into the activity, interruptions, and problems. The fourth or rank column is to assist you in evaluating whether the activity was important (I), unimportant (U), or had no relationship (N/A) to your daily objectives or a necessary condition (NC).

After 7 days of recording daily objectives, activities, comments, and importance ranks, conduct this analysis.

1. Study your daily objectives. Are they realistic (can they be accomplished in one day) and measurable (can I tell when I am finished)? “To scan and read [Chapter 3](#) of my history book” is a realistic and measurable objective. In contrast, “to study history” is a vague and immeasurable objective.
2. Rank your activities each day. Which activities were important to accomplishing your short-term objectives, unimportant to accomplishing your objectives, not applicable, but still important or a necessary condition activity? How much time fell into each of these categories?
3. Identify all activities that are travel, work, school, sleep, eating, and leisure. What percentage and how many hours each day were time wasters, as related to your daily objectives and longer-term goals?
4. Identify how you might eliminate or reduce these time-wasting activities.
5. Identify activities that you can control as well as those you cannot control but have to perform

(necessary conditions).

6. Classify and study the interruptions that occur each day. Why did they occur?
7. List UDEs related to your ability to manage your time.
8. Analyze these UDEs to determine their causes.
9. Take steps to eliminate the underlying cause of these time wasters.
10. Complete this exercise every six months.

It should be noted that while this exercise might seem tedious, it takes very little time to do. At the same time it adds enormous value to your understanding and ability to plan effectively for one of your most valuable assets—your time.

Developing a Detailed Implementation Plan to Accomplish Your Goals and Objectives

Time management is a skill that can benefit you throughout your professional and personal life. It provides you the opportunity to maintain a balance among your competing activities (personal, family, friends/community, work, and professional). Achieving your short- and long-term goals means developing a detailed implementation plan and being proactive in attacking your action plan. This action plan should be specific with respect to what, where, when, and how.

The first critical question, “What?” can be used in two ways to determine the significance of the relationship of an action to achieving your objective. First, what is the action that allows you to achieve your objective? This question

allows you to identify what action should appear in your “to-do” list. These significant actions must be accomplished to complete your daily objectives. Second, each action in your “to-do” list and actions you actually take during the day should be analyzed to determine what purpose they are serving. You should ask yourself, “How does this action help me accomplish my daily objectives?” If you are altering your daily plans, you should be aware of the sacrifice you are making. Is the new action that you are taking something that moves you toward today’s objectives? Is it worthy of taking your time from what you planned to do today? In many situations, you do not question interruptions to your plan.⁶ You accept them then and later reflect back on the waste of your time. Avoid interruptions and certainly question them. You might need to develop another strategy to accomplish your objectives. If you have many interruptions in studying, you might search for another strategy (of when and where) for studying.

The most effective way to control interruptions is to remove yourself from environments that permit them. It is responding to the second question, “Where?” Find yourself a quiet place to concentrate and get through your work with focused attention. You will be surprised at how much you can accomplish in a short period of time when you are able to devote 100 percent of your attention to one task at a time. The correct answer to the question of “Where?” can save time and effort.

To be successful at time management, the third question of “When?” must be answered repeatedly. The simplest way to improve your use of time is to use some type of daily planner (electronic or paper) to provide hourly, daily, weekly, and monthly calendars for recording your plans. At the beginning of each term, enter your important work and school activities, such as (for work) projects, reports, and meetings, and (for school) tests, term papers, projects,

football games, and parties on the monthly, weekly, and daily planning portions of the planner. Add your doctor, dentist, hairdresser, and club meeting appointments to your planner. The monthly and weekly overviews provide you an indication of what is coming up and times of peaks and valleys in your current workload. One approach many students have found useful in time management is to change their daily routine. Go to bed early on weekdays and get up early to study or exercise. You are fresh and alert and have few interruptions. Similarly, many white-collar professionals get to work before others so they have quiet time to get important tasks finished without interruptions.

The fourth question is, "How?" How are you going to accomplish your daily objectives? For example, how do you work or study best? For a project, do you first lay out all the tasks yourself or get the project team together and draft a project plan? In studying, for example, do you first skim a chapter; second, read it; third, go back and underline important items; and forth, review your underlined items? In an examination of used textbooks at a college bookstore, more than half (in some cases almost all) of the text within each chapter was highlighted. This suggests that most students skip the skim activity, and read and underline simultaneously. This undesirable approach was verified by taking a student poll in a number of college classes.

Operations Planning and Control Functions

Once you have determined your daily activities, You must focus on completing them. Focus translates into performing four interrelated functions which are required to plan and control your activities. These functions are priority planning, priority control, capacity planning, and capacity control. They are defined as:

1. *Priority planning: The process of determining the sequence of activities based on their relative importance.* What should be performed first, second, third? What should be set aside for another time? What should not be performed (errors of commission)?
2. *Capacity planning: The process of determining the time and resources required to perform a task (capacity required) and comparing them to your available time (capacity available).* In this planning, avoid multitasking; focus on one task at a time. Unless wait time is involved in an activity, focusing on one activity at a time usually reduces the time involved and improves the quality of the activity.
3. *Priority control: The process of executing the priority plan and making changes to the sequence based on current needs and conditions.* Interruptions happen; try to reduce the possibility of them occurring. Sometimes an interruption is important and you have to reprioritize your activities. Try to finish what you are doing before starting the new activity.
4. *Capacity control: The process of comparing your actual time and resources to perform a task to the capacity plan and making capacity adjustments to your work schedule based on your actual progress.* Most people underestimate the time required for an activity and then suffer the consequences. Track the accuracy of your time estimates so you can learn to estimate activity times better.

While these terms may seem foreign to you at first, you intuitively perform these functions throughout your daily activities. For example, you are planning your day and you want to drop off some clothes at the dry cleaner across town

after work, which ends at 5 PM. The cleaner closes at 6 PM. You have a 6:30 appointment for dinner with a friend. You compare your priorities and timing of the activities—priority planning. You estimate that you have one hour to get across town to the cleaners, and the round trip drive is 40 minutes—capacity planning. It is 5 PM, but instead of heading to the dry cleaners as planned, you are delayed because your boss asks for your help with a problem. You weigh the situation of heading to the cleaners versus helping your boss—priority control. You spend 40 minutes helping your boss and realize you cannot get to the cleaners before it closes. You mentally reschedule the cleaners for tomorrow—capacity control.

To be an effective manager, you must plan and execute these four functions with discipline. Some individuals are very good at performing these functions naturally; others are miserable. Most find planning and controlling their time a real problem, but do not seem to recognize the problem until it is too late. The key to effective time management is focus.

Steps to Improve Your Productivity

Formalizing the process of applying these four functions to your projects and daily activities helps improve your productivity. Follow the eight steps here.

1. Verbalize your objectives, supporting strategy, and performance measures.
2. List the activities that you must perform to attain your objective.
3. Prioritize the activities based on causal dependencies, start times, urgency, importance, ease of completion, distastefulness, efficiency, or some other basis.

Always do activities that move you closer to your goals in each facet/dimension of your life.

4. Estimate the resources (time, materials, and equipment) required.
 5. Compare the resources required to the resources available.
 6. Develop a simple plan for accomplishing your activities. Ensure that resources are available when needed.
 7. Focus on the activity at hand. Find a quiet place and time to perform critical tasks. Do it and move on to the next item on your "to-do" list.
Do not multitask on important tasks. Eliminate distractions. Try to buffer yourself against time wasters.
 8. As activities are accomplished, delayed, changed, or eliminated, adjust your list accordingly.
-

Now, each step will be discussed in more detail.

1. Verbalize your objectives, supporting strategy, and performance measures.
With your goals in mind, you must decide on your objectives, strategy, and how to measure your progress in attaining these objectives prior to defining how you will accomplish your objectives. This can be as simple as writing one sentence.
2. List the activities and projects (based on your strategy) you must perform to attain your objective.
QUESTION: How do you eat an elephant?
ANSWER: One bite at a time!

Accomplishing an objective is like eating an elephant. Identifying the activities required to accomplish an objective is critical. Projects such as complete the expansion plan, write a research paper for supply chain management, study for the history test, or clean the apartment have little meaning. Be more specific. What are the specific activities that must be performed to accomplish these projects?

List those activities that should be accomplished today. Do not start too many different and unrelated activities at one time. It is far better to focus your attention on a complete project and devote enough time to completing the project or a major activity of the project than to do a little of a lot of different activities simultaneously (multitasking). We discussed the problems of multitasking (briefly defined, it is moving back and forth across several different tasks at one time) in Chapters 3, 4 and 5 on project management.

3. Prioritize the activities based on causal dependencies (what must be done first, second, etc.), urgency, importance, timing, ease of completion, efficiency, distastefulness, or some other basis.

Do not over-prioritize. Some activities are urgent, while others are important. Some require starting at a specific time; some require large segments of time; and others require many small segments of time over a long period. Still others, while important, require little or no time. Remember your life goals and supporting objectives: Are you doing activities that move you toward these ends? The objective of prioritizing tasks is to remain flexible to respond to problems and opportunities.

4. Estimate the resources (time, materials, and equipment) required.

Estimate the resources required for each project to be completed. For example, to complete the expansion project report, you will have to set up an appointment with the contractor, you have to get estimates of equipment investment, you have to get permits, etc. To set up the project plan, you will have to set up a meeting with several different people. This part of the assignment will take approximately 4 hours, and the actual meeting to develop the project plan will take another 4 hours. Your capacity plan indicates 4 hours of work this week in setting up the meeting and the actual meeting scheduled for late next week.

5. Compare the resources required to the resources available.

Once you have a good idea of the resources required for an activity or project, you have to compare these estimates to the resources available. You might have 4 hours available this week and another 4 hours available late next week for the meeting. You probably need another 2 hours to prepare for the meeting early next week.

6. Develop a simple plan for accomplishing your activities.

Keep the plan simple (it can be a simple “to-do” list)! This plan entails identifying tentative dates and times for initiating each activity. When can you fit the 4 hours for contacting the project team? While you have 4 hours available tonight, most team members are working; therefore, you can’t use this time for setting up the project meeting. What high priority work can you accomplish this evening related to school? To complete your plan, you need to identify the next

priority item on your “to-do” list. You must ensure that resources are available when needed. The simple buffered “to-do” list discussed later in this chapter has proven highly effective for most students and managers.

7. Focus on the activity at hand. Find a quiet place and time to perform critical tasks.

Do it and move on to the next item on your “to-do” list. Do not multitask on important tasks. Eliminate distractions.

Focus, focus, focus! Clearly define the objective of the task and have all the materials needed to complete the task—then do it. Put yourself at a time and in a place that minimizes interruptions. Turn off your cell phone, the television, the radio, etc. This helps prevent multitasking.

8. As activities are accomplished, delayed, changed, and eliminated, adjust your list accordingly.

Check your plan (“to-do” list) frequently. When you complete an activity, mark it as complete. When you start or complete an unplanned activity, check the plan to see if you need to change or reprioritize activities. Murphy strikes! Murphy is the fictitious character who always disrupts plans. Murphy is alive and well and loves to create havoc with your plans. Any number and type of disruptions can wreck your plans. The objective in developing your plan is to recognize that Murphy will strike, and despite all good intentions, you cannot execute the plan exactly as established. Flexibility is the key—the ability to adjust your “to-do” list accordingly.

Using Buffer Management to Increase Your Effectiveness

Buffering is a critical activity that few people perform. Buffering your schedule can help you plan and control your daily activities in moving closer to your short-term objectives. Buffering protects the schedule from constant disruptions. Your “to-do” list incorporates the functions of priority and capacity planning and control. More importantly, it is based on time being a precious commodity and focuses on its effective use.

Let’s examine this buffering concept in detail. Your time is precious and your goal is to have a higher quality of life. You need to accomplish work, academic, social, family, service, and personal activities to accomplish this goal. You devote a certain number of hours each day, say, 10 hours, to accomplishing these activities. While you may work 10 hours a day, you may accomplish more or less than 10 hours of work—more by working faster than usual or less by having more interruptions than usual. You may estimate that you have 4 hours of work in calling project members to set up your expansion project, but you were able to complete the work in 2 hours. You consumed 4 hours of forecasted work; therefore, you need to have additional work available to start on or you will not take advantage of the 2-hour savings in actual time. If you never take advantage of time savings and you always suffer the consequences of delays and interruptions (Murphy), then you will always be behind your schedule.

The “buffer” is the amount of work (measured in time) you have planned and with you, ready to be worked on in case Murphy strikes and you cannot perform the next scheduled task. The objective of a buffer is to increase your effectiveness by planning to have the next activity (work) available to you when you complete your current activity. Additionally, the second highest priority activity should also be present in case an interruption prevents you from proceeding on the planned activity. In fact, a couple of hours

of high priority work should always be available for you to perform. Most people plan for only the current activity, and when Murphy strikes, they end up wasting the time allocated to that task because they did not plan effectively.

For capacity control purposes, the buffer should be divided into three regions (similar to the traffic light colors)—red (eminent tasks, the task that should be worked on from the present to the next few hours), yellow (lower priority tasks), and green (tasks to be performed later in the day). These regions are sometimes called region 1, 2, and 3. Region (red) contains the immediate activities to be performed; region 2 (yellow), the later priority activities; and region 3 (green), the last activities of those to be performed during the buffer period. Activities in region 1 are performed first, and as performed, activities from region 2 and 3 move up in priority and are performed as sequenced. If, for some reason, you cannot manage an activity in sequence, then move to the next activity in priority sequence and perform the skipped activity when you get a chance.

An example of the buffered “to-do” list is provided in [Fig. 38-6](#). First, you have to prioritize your activities (priority planning) and estimate their time duration (capacity planning). Notice several items (meetings, classes, appointments) in the list are time related. Next, you have to identify any requirements for accomplishing the activity (e.g., files, reports, books, notes, and meeting and class times). List the activities in priority sequence with any known requirements and the estimated capacity required (time) to complete the activity. The buffer has arbitrarily been set at approximately 15 hours of work (from leaving home in the morning to returning at night), with each of the three regions (red, yellow, and green) containing approximately 5 hours. You initially indicated spending about 10 hours per day working on your activities, but Monday is a particularly heavy day. You, in fact, will be away from your apartment for 15 hours. The buffer is set larger

than this time period to ensure that if Murphy strikes (e.g., the boss is busy and reschedules or your production meeting is canceled), you won't run out of work before you return to your apartment. Priority control (sticking to the plan) is accomplished as time progresses.

Region	Activity	Requirement	Estimated Capacity
Red	Uninterrupted office work	7-8 AM	1 hour
	Meeting to discuss expansion with boss	8-10 AM	2 hours
	Work on the monthly resource plan	10-12 AM	2 hours
Yellow	Lunch with Ann	12:10 Grill	1 hour
	Production meeting	1-3 PM	2 hours
	Work on the production meeting report	3-5 PM	2 hours
Green	Dinner/discussion with MBA project team	6:00 PM	1 hour
	Attend Supply Chain class	7:00-8:20 PM	1.33 hours
	Attend Managerial Accounting class	8:30-9:50 PM	1.33 hours

FIGURE 38-6 An example of a buffered “to-do” list.

Suppose your boss calls and reschedules your 8 AM meeting for 10 AM. You can check your “to-do” list, move your task of working on the month resource plan to 8 AM, and then go to the meeting with your boss. As time progresses, you also have to make adjustments to your capacity estimates—you estimated 2 hours to complete your work on the resource plan, but suppose after 2 hours you still have 20 minutes of work left. You decide to reschedule your lunch for 12:30 and finish the resource plan prior to lunch. You call Ann to reschedule and ask her to order your lunch when she places her order.

Notice that as you performed activities, you progressed down the list with activities in region 1 (red), which has the highest priority for your time. If you could not accomplish an

activity in order, you moved on to the next highest activity in which you could meet the requirement.

Buffer Management is a simple approach to increasing your effectiveness because it provides a time buffer of activities at your disposal. Be assured, Murphy will always strike, so you must be prepared. The key is always to have the next highest priority activities available to be worked on just in case something does not go as planned. Failure to buffer your work results in unplanned idle time, working on unimportant activities, and having the wrong items to complete an activity. At the end of the day, you should examine your buffer list to plan your next day's activities. If you did not complete your list of activities, remember that you have accomplished the most important ones. You should move any incomplete activities to the next day and prioritize them based on planned activities listed in your daily planner for the next day. An additional purpose of BM is to identify the cause of disruptions to your schedule. An analysis of the causes of disruptions should be performed to identify which causes (maybe, your cell phone, or watching TV while you study) must be addressed to improve overall performance.

Don't be discouraged if you only accomplish half of the activities listed in your buffered "to-do" list. You have to learn to estimate your capacity for completing activities, and more importantly, you have to learn how to control interruptions. Interruptions are a fact of life. Some are uncontrollable and disrupt your schedule totally. Planning to finish your activities ahead of time is actually a method of buffering these activities against interruptions. Suppose your MBA team project is due next Friday (recall you have a busy Friday planned already), you discuss with your team tonight the possibility of completing and turning in the project by Wednesday evening. This gives you a two-day completion buffer for your team report.

Several other guidelines for managing your time effectively are provided here.

- Set aside some quiet time for thinking and planning.
- Identify your creativity or energy cycle. Protect creative hours from interruptions.
- Schedule (sequence) the entire day, not just appointments.
- Have specific, realistic, attainable, and measurable activities to be completed in your time buffer.
- Eliminate or screen interruptions. (Cut off your cell; put a “do not disturb” sign on the door.)
- Find a quiet, isolated place to work on those critical projects.
- Group errands—getting items from grocery store, bookstore, and library. Travel times may outweigh activity times.
- Gain control—most individuals plan well, but they fail to execute.
- Establish daily, weekly, monthly, and annual objectives linked to your life goals.
- Set measurable objectives.
- Measure progress toward your objectives.
- Recognize that work requires focus, concentration, motivation, and time. If you cannot apply the first three requirements, then significantly more time is required.
- Eliminate multitasking as much as possible. Start and finish a task; the preparation for completing a task in many cases exceeds the time to complete the task. Starting over requires repetition of this timely preparation.
- Always have additional high priority work available to substitute where Murphy strikes and the completing of a scheduled task is delayed or when you finish a task ahead of time.

- Allow and schedule time for high priority activities that support your short-term objectives and goals in each dimension of your life—self, family, friends/community, work, and professional. A balance is required for effectiveness, satisfaction, and productivity.
- Reward yourself. Plan a rewarding activity for the completion of a difficult activity or a successful week. The reward might be as simple as a night out or a weekend trip with friends. Push to finish tasks before leaving so that the reward is meaningful.

Using the Thought Processes to Achieve Life Goals

This section presents one of the first applications of Goldratt's TP to achieve one's life goals. In 1992, after attending a workshop on the TP taught by Dr. Goldratt, I (Jim Cox) came back to campus with the idea of teaching these tools to my Advanced Operations Management class. I wanted to attack what I thought was one of the biggest barriers for students—personal productivity. I felt that the tools were extremely powerful in providing students a framework for logical analysis of any problem. What better area to study than keeping the many challenges of student life in balance?

After I made the assignment of using the TP to improve personal productivity, Sheila Taormina came to my office and asked not to have to do the assignment. Sheila was an exceptional student—swimming several hours every day, maintaining a near 4.0 grade point average, and active in numerous student organizations as an officer. Instead of analyzing her personal productivity, Sheila wanted to analyze her swimming. I really didn't have much hope in convincing her that she needed to study her productivity.

However, I was somewhat perplexed at the time. I can't even float, so how was I going to help her analyze her swimming? She was an All-American swimmer! I verbalized my concerns, but agreed, knowing that she was not trying to get out of work, but really wanted to learn something that might help her swimming career. This story is the result of that "personal productivity" project.

Sheila's Story

By Sheila Taormina

In November of 1996, I opened the doors to my very first house. I certainly did not have much furniture to fill the rooms, but I did have a van-load of boxes containing mostly knick-knacks that I had collected throughout my college years at the University of Georgia. As I came across the box filled with my old college papers (the ones I saved that I thought would be fun to look over in future years), I remembered a project I had done, and I prayed it would be in the box.

There it was . . . MAN 577, Spring 1992, Dr. James Cox, Personal Productivity Analysis. A flood of emotions came over me because I knew that the work that went into this project was the catalyst that led me to a dream come true. As I read the pages, I relived every feeling that I had in 1992—a time in my life that was filled with questions and anxiety, but most of all a fearful kind of hope. Now, in order to understand the rest of this true story, you will need some background information that will take you to the point of the spring of 1992 when I wrote my paper for Dr. Cox.

I have been on a swim team since the age of six, and in 1988 and 1992 I qualified to compete in the Olympic Swimming Trials. I was 18 years old in 1988 and 22 years

old in 1992, which typically are the peak years of swimming for females; however, I missed making the Olympic team both times. I was not disappointed in failing to make the team though, because I never expected to make it. After all, I believed that the people who make the Olympics are a level above all of us average people . . . they have some special talents.

My plans were to retire from the sport of swimming after the 1992 trials, but when my friend made the team to Barcelona that year, it was as if a light bulb went off in my head. I suddenly realized that I had been defeating myself all of these years before even stepping up to the starting blocks. My friend was not superhuman; he had no special talents! I immediately had a desire to make an attempt at the 1996 team.

My problem was that I was already 23 years old, so if I stayed around for four more years I would be 27—a dinosaur in the world of women’s swimming. In addition, I was finished with my collegiate athletic eligibility, so I would have to support myself financially. My biggest dilemma was figuring out a way to drop three seconds in my 200-meter freestyle in order to get a fast enough time that I thought would have a reasonable shot at making the Olympic team. If you know anything about competitive swimming, then you know that three seconds is quite a bit of time to drop.

I was not guaranteed that dropping three seconds would be fast enough anyway because what I was doing was very similar to a sales forecast for a business. I looked at the history of women’s swimming and figured that a 2:00 (two minutes even) was a safe bet, but as every business manager knows, forecasts are not always reliable!

I suppose, however, that my philosophy amidst all of this fear and questioning was, “I just do not want to look back 50 years from now and wonder ‘what if . . .?’” The one thing I knew for sure was that I needed to return to Georgia and finish my last quarter in order to get my bachelors degree

(specializing in Production/Operations Management). I enrolled in my final three classes, one of those being Management 577 with Dr. Cox. I had taken a few production management courses with Dr. Cox already, and he was always understanding of my swimming schedule.

We started the quarter learning the fundamental productivity tools: Current and Future Reality Trees, Evaporating Clouds, Transition Trees, and Prerequisite Trees. Our first assignment was to apply these tools to a situation in our personal lives, a personal productivity analysis. Of course, the first and only thought to come to my mind was, "How do I get faster to have a chance at making the '96 teams? What have I done incorrectly in the past, and how do I change that?"

This assignment was in-depth, and I loved every minute of it. I knew that it would help me identify the track I needed to take for reaching my goal. I spent many days after class asking Dr. Cox for help in preparing my tools correctly. He helped to point out the missing links in my *thought* process. As my Current Reality Tree was beginning to come together, I started building confidence in myself. The most amazing realization was that my core problems were not larger than life!

The following pages are the actual paper that I wrote in 1992.

Personal Productivity

Sheila Taormina
Man 577, Dr. Cox

Spring, 1992

Here was the scene: There was a fireworks display in the Natatorium while the National Anthem played in the background. A huge American flag dropped from the ceiling

and the people inside erupted as the 41 members of the United States Swimming Olympic Team paraded around the pool. It was a send-off for the swimmers who will be going to Barcelona.

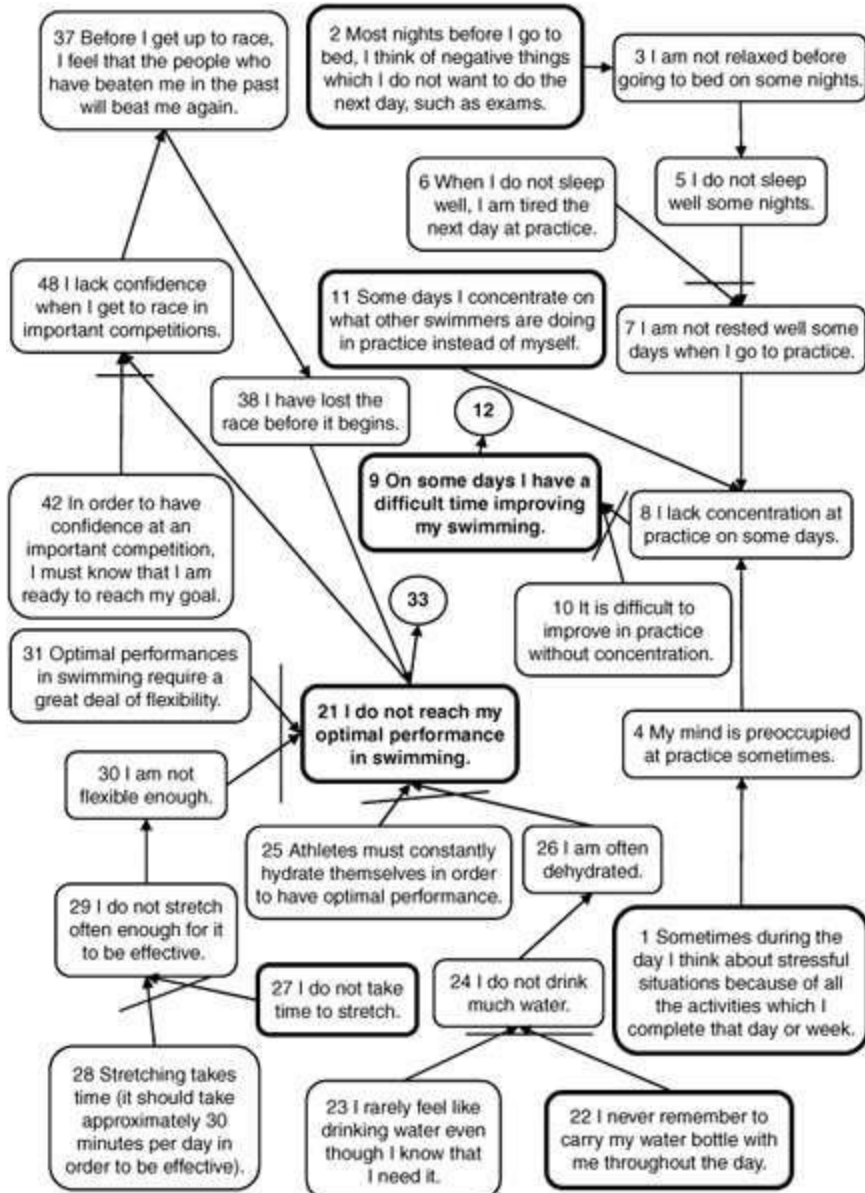
I thought that I would retire from swimming after the Olympic Trials in March, 1992. Even though I tried to convince myself that I could make the team to Barcelona, deep inside I had no confidence. When the trials were over, I could not bear the thought that I had just posted some personal best times in two of my events and was going to quit swimming while I had the opportunity to learn more and improve. I have found it interesting to complete a Thought Process Analysis for a personal problem, and it helps me to understand why it is essential that business entities should look deeply into the problems which face them. I always believed that there was no need for a business to constantly strive to be the leader in the industry, because, as long as a profit was made, then what is the big deal about claiming the number one position? This personal analysis has made me realize why businesses compete on a continuous basis. One competitive disadvantage can be the difference between reaching a goal or not, and when a few of those disadvantages are put together, it is sometimes amazing that a company or individual is still in the game at all.

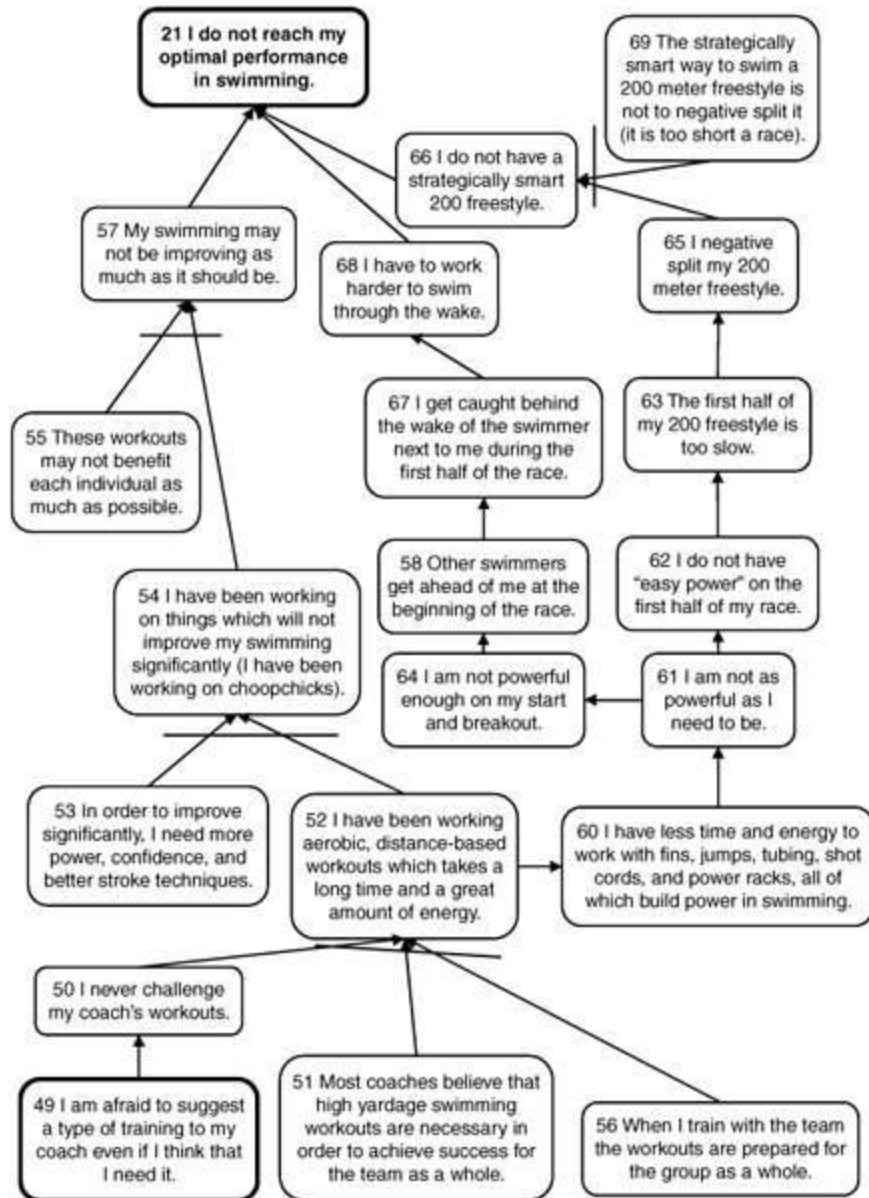
I think that I stayed in swimming after the trials because I was still learning in each practice how to improve, and I wanted to give myself the chance to use what I learned. I am 23 years old now, and although most female swimmers peak from the ages 18-20, I have been able to break the tradition through a process of continuous improvement. I also believe that there is another reason why I am still swimming, which is the fact that I am enjoying it so much right now, and I am finally seeing the results of many years of hard work pay off.

My hard work is paying off in more ways than one. I believe that the countless yards/meters I trained during my

high school and college years have formed an aerobic base on which I can rely. Now I need to refocus my energy on improving in the areas which I have not worked on a great deal in the past eight years. Before I go into a detailed analysis, I would like to direct any readers who are not familiar with swimming jargon to refer to the Appendix entitled, "Definitions." (Not included) Also, there is one other clarification I need to state: in manufacturing, efficiencies can be considered a negative measure unless used at the constraint; however, I speak of efficiency in swimming as a positive measure. When I refer to it in this paper, I am speaking in terms of technique, such as streamlining in the water, hand pitch and hand entry in the water, elbow position, head position, and shoulder roll. An efficient stroke allows the swimmer to have "easy power."

It happens that my technique in swimming is average, but I have found it to be a negative effect. Other negative effects, some of which I could identify off the top of my head and others which I never thought of until they appeared in the Current Reality Tree ([Fig. 38-7](#)), include: (1) I am not as powerful as other female swimmers, (2) I do not have adequate flexibility, (3) I am dehydrated often, (4) I do not get a good night sleep often, (5) I am afraid to race the top swimmers in the world, because I do not think that I can win, and (6) I am not ranked as high in the world as I would like to be or have the capability to be.





2. I am stressed during the day to accomplish many tasks,
 3. I pay too much attention to what other swimmers do in practice when I should be paying attention to my own swimming,
 4. I push beyond my physical limit in practice sometimes,
 5. I never remember to carry a water bottle,
 6. I do not take time to stretch, and
 7. I am afraid to suggest a type of training to my coach even when I think that I need it. (Paraphrased from tree where the core problems are shown in bold.)
-

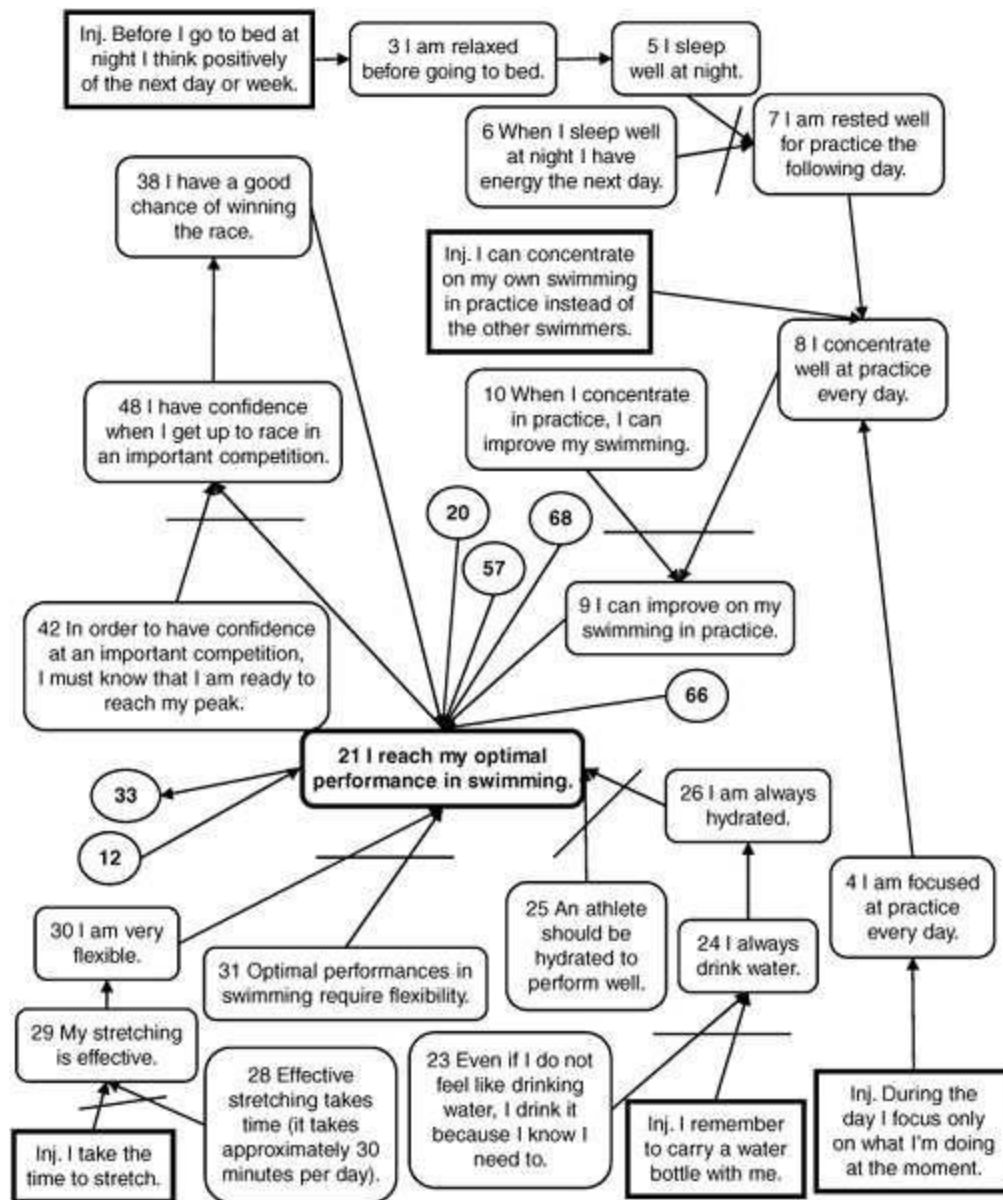
The connection among the core problems and negative effects is as follows: If I think of things that worry me before going to bed, then I am not relaxed when I go to bed. If I am not relaxed before bed, then I do not get a good night's sleep and am not rested well for practice the next day. The second core problem of lacking focus during the day causes a lack of focus and concentration in practice. If I do not concentrate, then my technique is poor. In swimming, it is difficult to maintain the correct technique when you are tired or not focused on your stroke. The third problem of paying too much attention to other swimmers and not enough to my own swimming causes me to lose the necessary concentration on my stroke. The fourth core problem of pushing beyond a certain training limit causes two negative effects to take different paths. The first part is that I have a lack of energy when my body is broken down, and my stroke efficiency once again suffers. The second path is that my body takes a long time to recover when I

break it down too far. If my body takes a long time to recover, then I may not be able to perform well for the next practice. In fact, I have been so broken down before that I could not keep up with the team in practice for three weeks. I finally took four days away from any type of training and was able to recover. The next two core problems of lack of stretching and lack of water bottles both lead to a less than optimal performance in a competition. Stretching is essential for competitive swimmers as it is for most athletes, and each athlete should stretch for 20-30 minutes per day. When I forget my water bottles every day, then I am dehydrated, which is dangerous for training. The negative effect is the same as before . . . I do not reach my optimal training performance nor do I reach my optimal competition performance.

The final core problem is the most difficult to overcome in my opinion, and it leads to many negative effects of my current situation. I am afraid to suggest a different type of training to my coaches because I do not want to show disrespect for their scheduled workouts. However, I feel as though their workouts were exactly what I needed up until this stage of my swimming career. When I do not communicate with my coaches about practice, then I work on the wrong things. If I work on the wrong things, then I am not improving in my swimming. If I do not improve, then I do not reach my optimal performance. Furthermore, if I am working on the wrong things (such as aerobic base), then I lack the necessary power I need for the 200 meter freestyle (the event which I feel I have the most potential). If I lack the necessary power for the 200 free, then the first half of my race is going to be slow. When I am too slow in the front half of my race, then I get caught behind the other swimmers' wake and have a difficult time passing them in the second half of the race.

These negative effects can be eliminated if I could effectively implement a plan to change my core problems

into positive actions. Before developing an implementation plan, I have constructed a Future Reality Tree ([Fig. 38-8](#)) to see the effects of making the core problems into positive actions. The ultimate result is a reverse of the negative effects in the Current Reality Tree. One comment I must make here is that the Future Reality Tree indicates that I will reach my optimal performance. I cannot be guaranteed that my optimal potential performance will take me to my goal of improving my world ranking to the top eight. Furthermore, I must be careful that inertia does not set in. I could focus too much on training for a power base and completely ignore my aerobic base. I am aware that my aerobic base will be lost if I neglect it; therefore, my training will always include adequate work in this area.



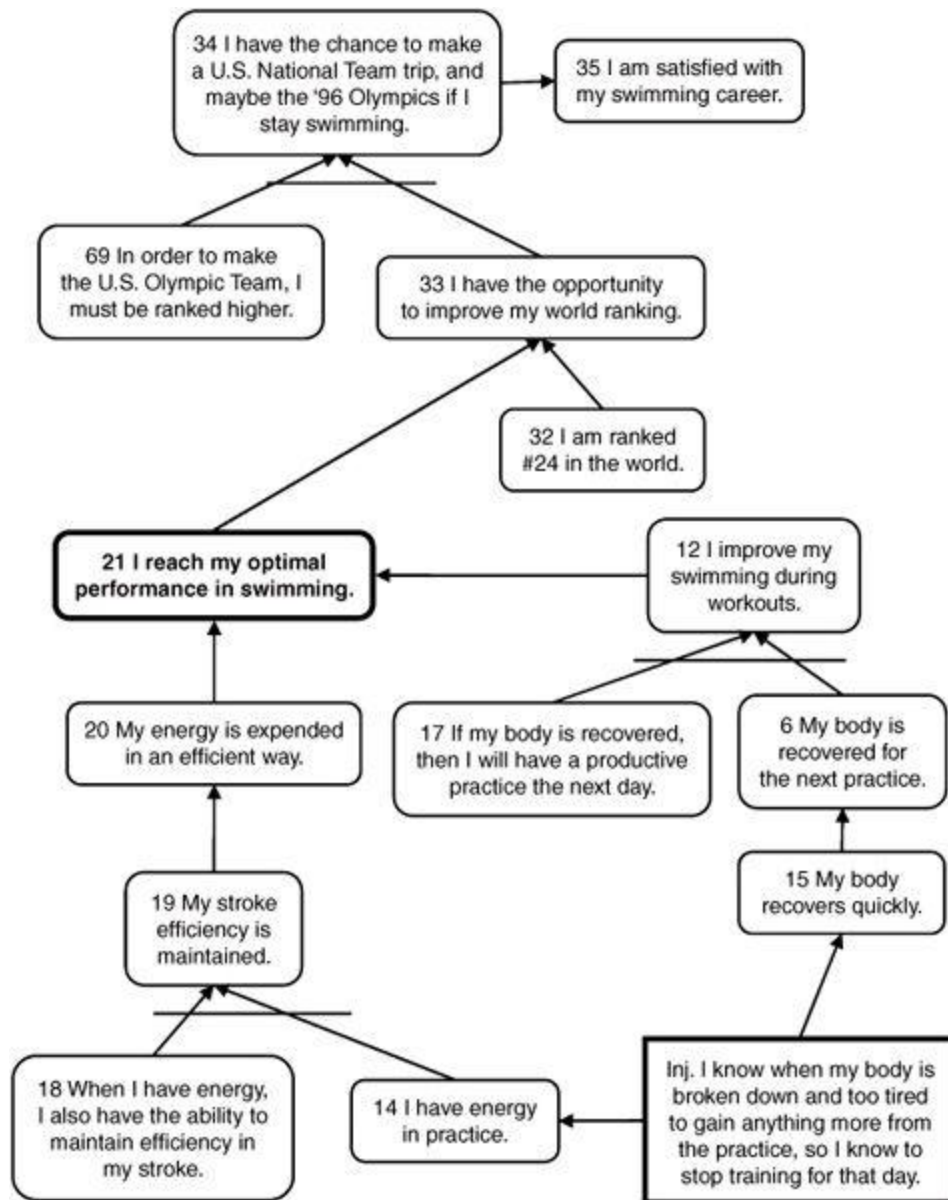
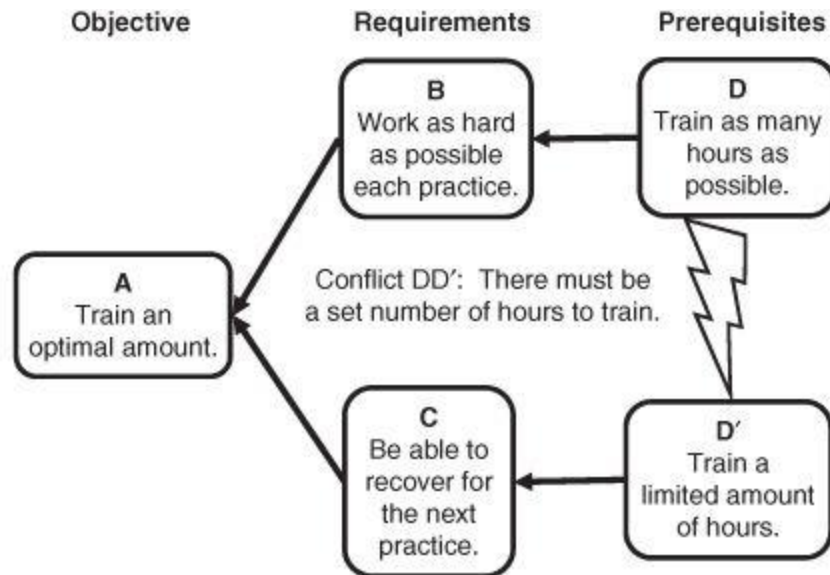


FIGURE 38-8 Future reality tree of Sheila's swimming.

The Evaporating Clouds in Figs. 38-9 through 38-12 mainly challenge the assumptions by which coaches and swimmers have always lived. The conditioning of an athlete includes many different objectives, including physical and mental training. A plan of action is necessary in order to measure how effective the training schedule is during the different times of the season. The key to success lies in developing an intelligent plan of action which breaks away from the old

paradigm that the more yards/meters a swimmer does, the better that swimmer will be. A coach and swimmer must develop the plan together in order to have input from both sides.

When the assumptions of the clouds are understood and managed in a beneficial way, I must begin to plan the actions to take in following through with my goals. The Transition Tree and Implementation Plan at the end of the paper outline the steps to take to achieve my plan. A few obstacles which I may encounter are identified in the Prerequisite Tree ([Fig. 38-13](#)), but I have developed another set of objectives to overcome those obstacles. I feel that everything in my plan is feasible and will help me to reach my goals. See the I/O map in [Fig. 38-14](#). The interesting part of all of this is that I will be willing to bet that the negative actions which I have been doing require more energy than do the positive actions. I have not reached my peak yet, and I will keep searching for ways to climb up the ladder of world rankings.



Assumptions -

AB - Assume that there are no negative returns in swimming.

Injection: There is a point at which your body can break down so much that it takes many days to recover.

AC - Assume that the next workout is scheduled to be hard.

Injection: The next workout may be scheduled to be recovery.

BD - Assume that working hard means many hours of swimming.

Injection: An hour anaerobic workout takes about half the time of an aerobic workout and is actually more difficult.

CD' - Assume that recovery requires as little swimming as possible.

Injection: A swimmer can recover by swimming at a low intensity and still make improvements in his/her stroke technique.

This cloud is broken at every link due to the injections stated after each assumption.

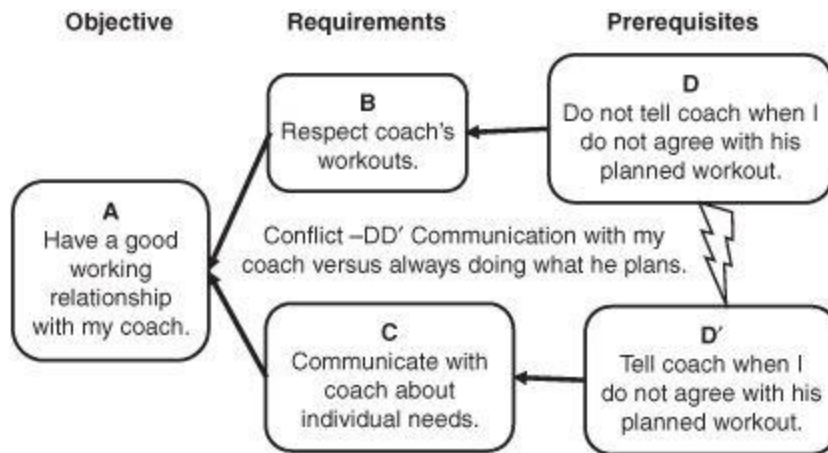
FIGURE 38-9 EC of training hours dilemma.

Sheila's Epilogue

I took this paper very seriously. I implemented the solutions, but it was not without challenges. I followed through with addressing every issue at some point before 1996. Some areas take more effort to correct than others do. For instance, drinking water in order to stay hydrated was much easier to implement than the visualization techniques and

positive thinking. I did not wake up one day a positive thinker! The process developed over time with practice.

The benefit that individuals will see the most from doing a productivity analysis is the identification of core problems and a logical way to find a win-win solution. Implementation depends on the conviction of the individual. I was determined to follow through with every effort in order to realize my positive effects. I moved home to Michigan in 1994 and trained with the coach I had been with since age nine, Greg Phil. Greg was not my first choice of a coach because the pool where he trains his swimmers is not a first-class facility with high-tech equipment. An interesting side story is that the team that I wanted to join in Colorado would not invest their time in me because they did not believe that I could make the team. Greg believed in my plan, and together we added what was needed that I had not yet identified. We even set a benchmark. It was simple: if I did not swim a 2:02 or better by the summer nationals in 1995, then we should put swimming behind us. Thank goodness, I swam a 2:02 that summer!



Assumptions -

AB - Assume communication leads to a good working relationship.

Injection: I can be angry with my coach at times and still respect him.

AC - Assume communication leads to a good working relationship.

Injection: Poor communication is not good for a relationship.

BD - Assume that staying quiet in practice shows respect.

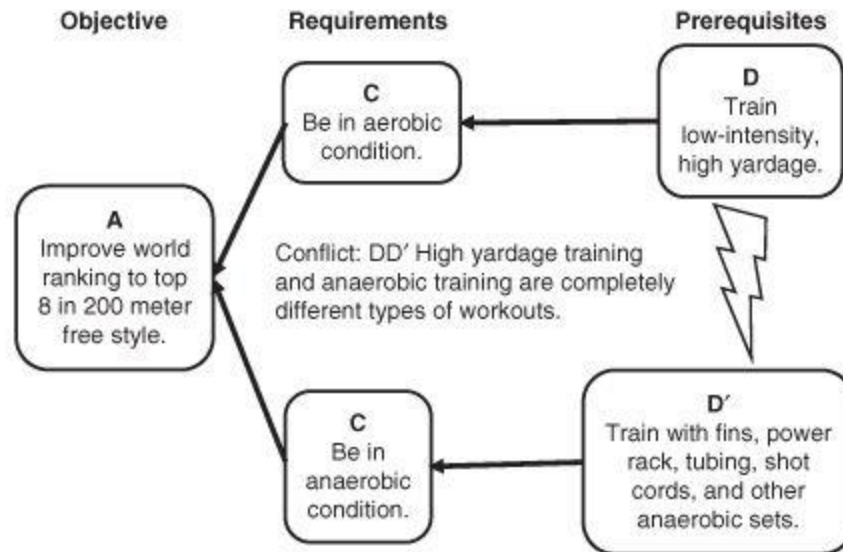
Injection: If I am not agreeing with his scheduled workout, then I get angry and show less respect.

CD' - Assume that communicating my concerns will change how my coach schedules practice.

Injection: My coach may believe firmly that I need what he scheduled.

All of these assumptions can be broken, but I have not attempted to fully communicate with my coach; therefore, I am not sure which ones apply to me and which ones do not. I will find out after I attempt the steps in my implementation plan.

FIGURE 38-10 EC of the athlete's and coach's communication dilemma.



Assumptions -

AB - Assume that I could be in better aerobic condition, which will improve my swimming.

Injection: My aerobic condition presently is such that any improvements will be minor.

AC - Assume that more anaerobic training will improve my swimming a great amount.

Injection: This is broken only if my anaerobic training gives me minimal returns.

BD - Assume that I am not in aerobic condition.

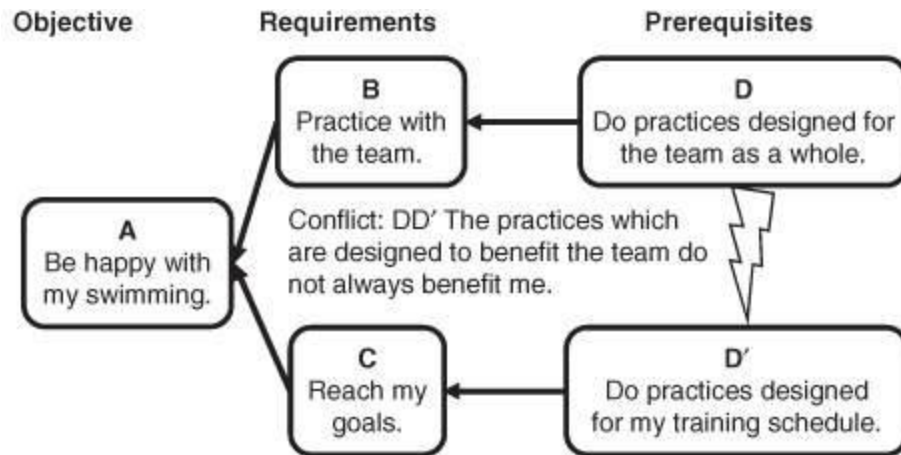
Injection: I have done aerobic training throughout my life.

CD' - Assume that fins, tubing, etc. develop anaerobic conditioning and power.

Injection: This does not work unless the swimmer uses them with 100% effort.

The cloud is broken at AB and BD, because I have a very high aerobic capacity, and any improvements will be so minimal that my times and ranking will not improve.

FIGURE 38-11 EC for the type of training dilemma.



Assumptions -

AB - Assume that I enjoy swimming with the team.

Injection: Not if the team is one in which the people always complain.

AC - Assume that I am satisfied with swimming if I reach my performance goals.

Injection: Not if I sacrificed something else that was important to me.

BD - Assume that when I practice with the team, I have to do their workout.

Injection: Not if I communicate with the coach and swim in a different lane where the swimmers are doing the same workout that I need.

CD' - Assume that I will reach my goals if I do the workouts designed for my training schedule.

Injection: Not if my mental training is so poor that it offsets the benefits of my physical training.

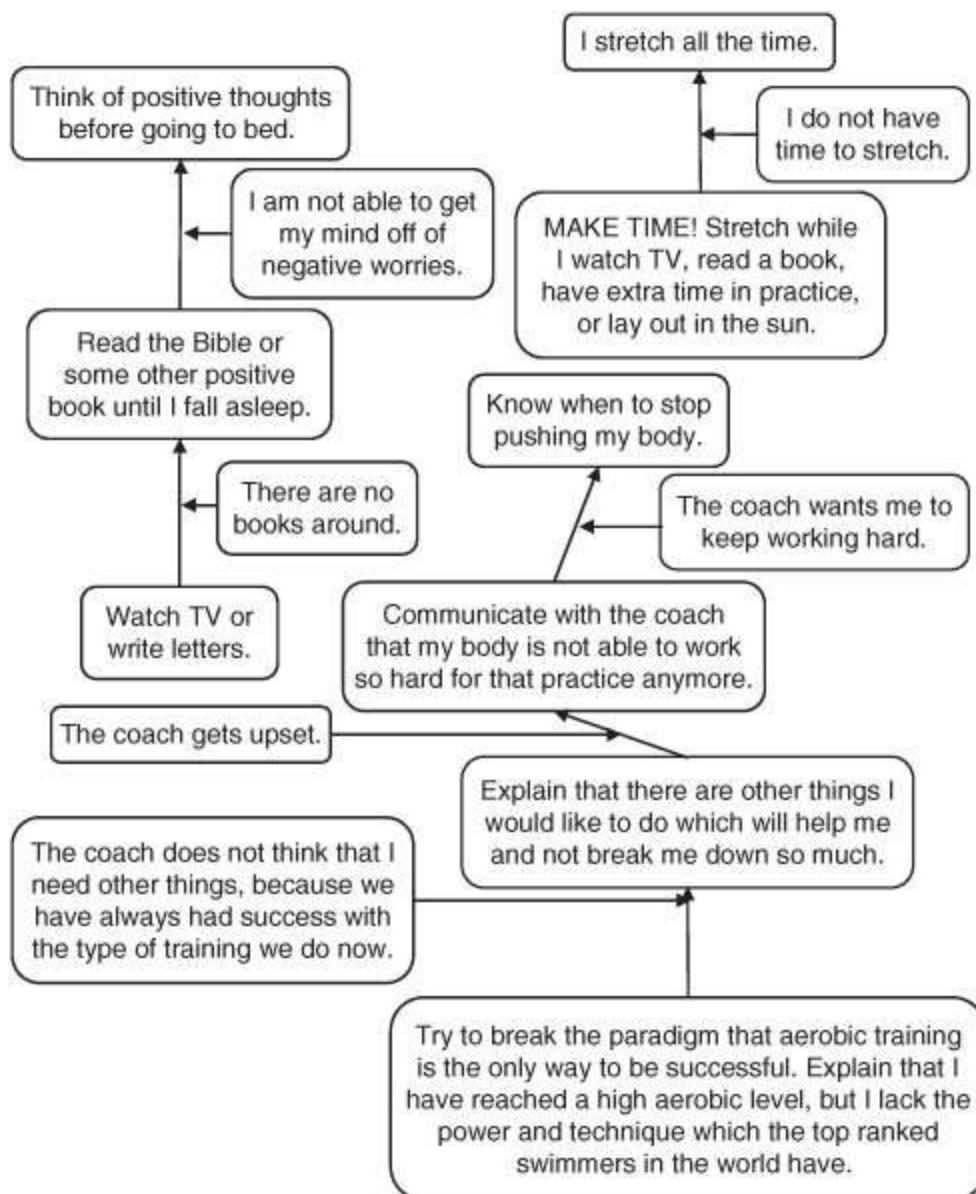
The cloud can be broken at any link for my specific situation; however, it is most likely broken with the injections at BD and CD'.

FIGURE 38-12 EC of the training schedule dilemma.

I will never forget the day when my full plan came together. Greg and I went out to breakfast the week before the 1996 trials because I was getting nervous about swimming. He and I had developed an action plan two years prior for how we were going to swim a 2:00 in the 200 freestyle (much of the action plan was derived from my MAN 577 paper). He pulled out that plan while we were eating. He read down the list, asking me if I had kept my promise throughout the past two years on every item on that list. I could answer “yes” to everything. He looked at me and said, “I don’t know if you are going to make the Olympic team next week, but I do know that you are going to have the

best swim of your life. You have already succeeded because you did everything in your power to give yourself a chance.”

He was right! The weight of the world was lifted off my shoulders at that moment. We drove to Indianapolis and I had the best swim of my life. I made the Olympic team by a fraction of a second. Then, beyond my wildest dreams, in Atlanta, in the 4 × 200 free relay, on which I swam the third leg, won the gold medal in an Olympic and American record time. My preliminary time in Atlanta: 2:00.57, my target time!



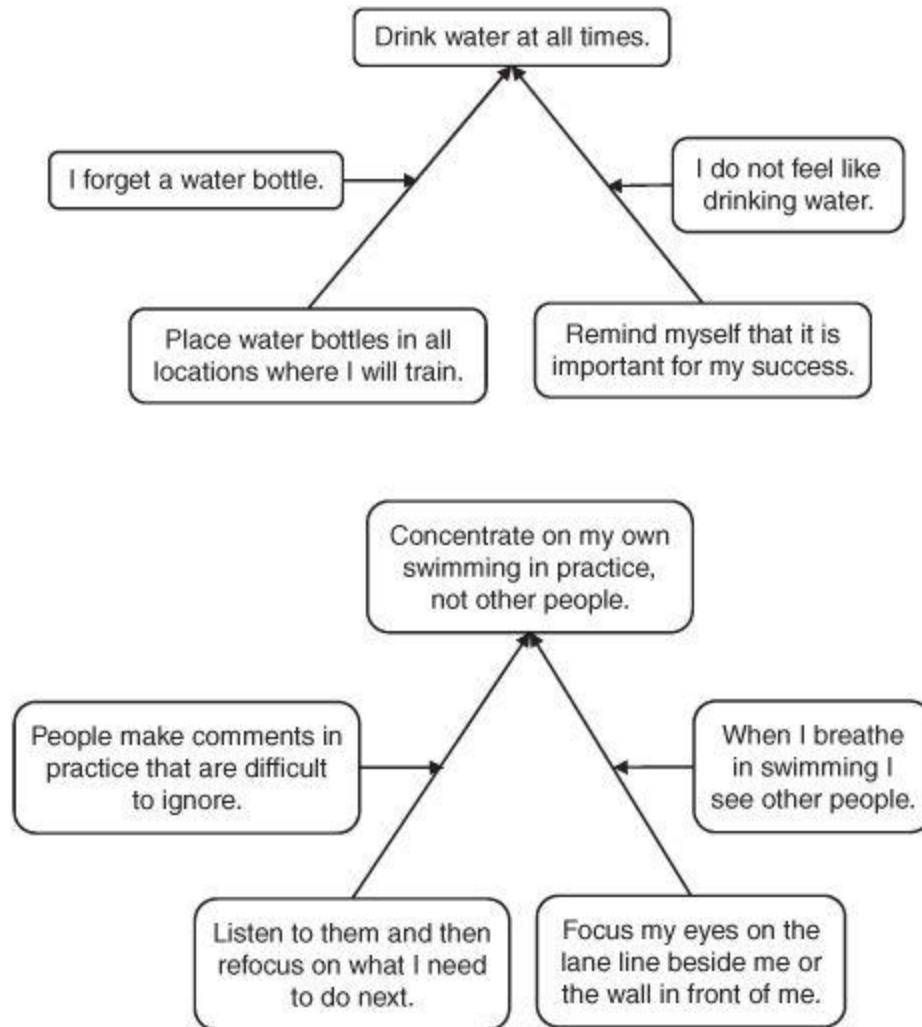


FIGURE 38-13 Prerequisite Trees of implementing Sheila's swimming program.

Upon returning from the 1996 Olympics, I went back to my job in the auto industry as a quality representative. About nine months later, I saw the opportunity to start my own speaking business, and for 2 years I traveled the country giving swim clinics and motivational talks. Finding myself terribly out of shape in 1998, I decided to try a local triathlon (in Ann Arbor, Michigan). The race director saw me and approached me to say that he thought I had some potential in the sport and that he would be happy to give some guidance if I wanted to pursue it further. I initially

turned him down, but then I decided to join his running group to stay in shape. It turned into a fun hobby, and before I knew it, I was flying to Africa to race (first pro race in March 1999). I made the Olympic team in 2000, and placed sixth in Sydney. The distance is in kilometers: 1.5 K swim (approximately 1 mile), 40 K bike (approximately 24.8 miles), and 10 K run (6.2 miles). That is the Olympic distance. I am staying in the sport as long as my body agrees and as long as I see that it is the direction God has planned. The 2004 Olympics are in my thoughts, but nothing is definite. Perhaps I will try the Ironman distance one day (2.5-mile swim, 112-mile bike, and 26.2-mile run).

Our Epilogue on Sheila

Sheila certainly exceeded her personal goal of making the Olympic Swim Team as an alternate. She credits the use of the TP for her success.⁷ We recognize that the Thinking Processes are quite useful for helping one attain or exceed one's life goals. Since Sheila's experience, we have seen numerous students use the Thinking Processes to improve their grades, get that special job, lose a significant amount of weight, get in great physical shape, and so on. It is not uncommon to see them go beyond what they were originally striving for.

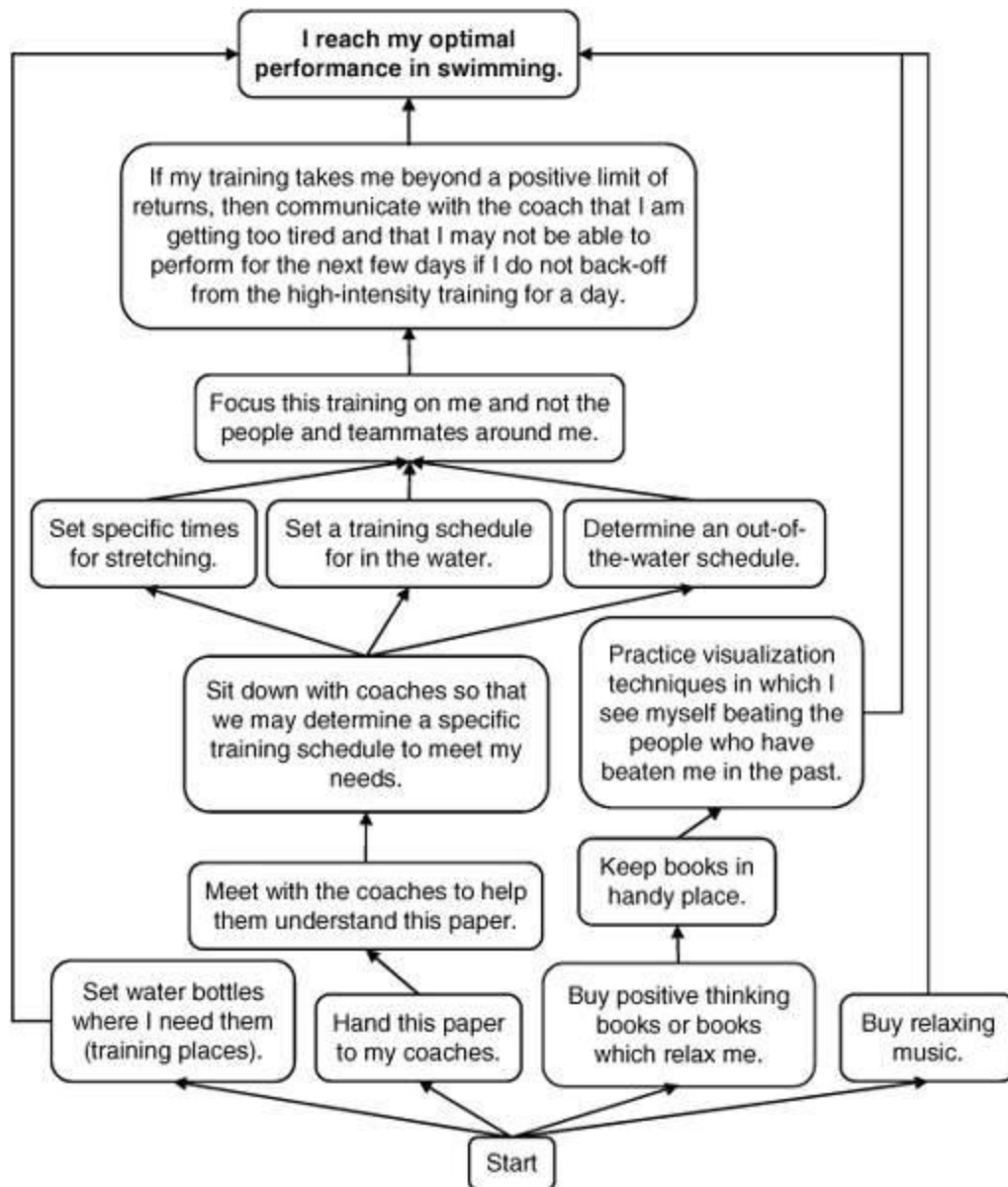


FIGURE 38-14 Input-output Map of part of Sheila's implementation plan.

Summary

Achieving one's life goals is generally only a dream for most people. Without having goals, strategies, supporting objectives, actions, and a measurement system indicating

your progress toward your objectives and goals, it is virtually impossible to achieve them. Developing a detailed life plan in your life facets is extremely time consuming but quite rewarding. However, obstacles block the achievement of your goals. Conflict is a major obstacle draining the motivation, concentration, and energy, and thus increasing the time and effort required to complete even the simplest tasks. A conflict in one facet of your life reduces the ability to focus on the task. Identifying and resolving these conflicts with a win-win solution is the key to gaining and maintaining the motivation, concentration, effort, and energy, and thus reducing the time required for the tasks at hand. Many of these conflicts, particularly with family members and associates at work, are actually chronic conflicts periodically surfacing repeatedly in different situations. You must recognize these situations as emanating from the chronic conflict and devise win-win solutions to these specific dilemmas. Once this approach (construct the EC, provide your assumptions, and best guess at the other side's assumptions) is taken, better short- and long-term solutions are surfaced. This will certainly make your life more pleasant.

Both the classic student and the classic burnout ECs and their assumptions should be studied carefully. Do you exist in one of these ECs? Does a child (or children) exist in the college dilemma? There are simple and effective solutions to both dilemmas but they have to be constructed and understood by the person with the problem. For example, many students have decided to identify the facets of their life (school, work, personal, family, friends, etc.) and the dimensions in each facet (personal might be divided into physical, spiritual, and mental, for example). They then identify the necessary conditions, goals, supporting objectives, and measures in each dimension. Once identified, they then determine the actions to be completed each day to help them accomplish their supporting

objectives. Some turn their day upside down (because of the interruptions in their normal day): they wake up at 4 or 5 AM and study until classes start, they attend classes, they socialize with friends in activities such as exercising, and they go to bed around 8 or 9 PM. They set up a reward system for a productive week by partying on the weekend, visiting family and friends, etc. They try to focus on one task at a time and complete that one as much as possible before moving to the next task. The same routine applies to business professionals—many get to the office early so that they have quiet time to work before the chaos begins. In this manner, they complete 1 hour of focused work, which may have taken 4 hours of normal office time.

Many devices have been suggested to increase your productivity; the best one we have seen is the daily “to-do” list of the top 10 items to be completed. The buffered “to-do” list works extremely well for students and also for some professionals. Murphy is going to strike; the buffered to-do list allows you to be prepared. Identifying and minimizing the causes of Murphy increases personal productivity even more.

References

- Covington, J. 2009. *Enterprise Fitness*. Mustang, OK: Tate Publishing.
- Cox III, J. F., Blackstone Jr., J. H., and Schleier Jr., J. G. 2003. *Managing Operations: A Focus on Excellence*. Great Barrington, MA: North River Press.
- Cox III, J. F., Mabin, V. J., and Davies, J. 2005. “A case of personal productivity: Illustrating methodological

developments in TOC,” *Human Systems Management* 24:39–65.

Goldratt, E. and Winter, L. 1996. The application of TOC for the individual. Presentation (video) presented at the Jonah Upgrade Workshop, Washington, DC: The Goldratt Institute, March 18–21.

Goldratt, E. M. 1993. “What is the Theory of Constraints?” APICS—The Performance Advantage. Reprinted in *Selected Readings in Constraints Management*. Falls Church: APICS, 1996.

Goldratt, E. M. 1994. *It's Not Luck*. Great Barrington, MS: North River Press.

Goldratt, E. M. 1995. *Management Skills Workshop*. Workbooks 1–5. New Haven, CT: Avraham Y. Goldratt Institute.

Sullivan, T. T., Reid, R. A. and Cartier, B. 2007. *TOCICO Dictionary*. <http://www.tocico.org/?page=dictionary>

About the Authors

JAMES F. COX III, PhD, CFPIM, CIRM, holds TOCICO certifications in Production and Supply Chain, Performance Measurement, Critical Chain, Strategy and Tactics, and Thinking Processes. He is a JONAH's JONAH, Professor Emeritus, and was the Robert O. Arnold Professor of Business in the Terry College of Business at the University of Georgia. He has conducted numerous academic and practitioner Theory of Constraints workshops and programs on performance measurement, production, supply chains, management skills, project management, and the thinking processes.

Dr. Cox's research has centered on the Theory of Constraints for over 25 years. He has authored or co-authored three books on TOC and almost 100 peer reviewed articles. He was the co-editor of the *APICS Dictionary*, 7th, 8th, 9th, 10th, and 11th editions, and an invited contributor on the topic of *Constraints Management to the Production and Inventory Management Handbook*.

Dr. Cox has been a member of APICS for over 30 years, holding chapter, regional, and national offices. He served on the APICS Board of Directors for four years with two years as VP of Education—Research and served on the APICS Educational and Research Foundation Board of Directors for nine years with four years as President. He was a founding member and elected to the founding Board of Directors of the Theory of Constraints International Certification Organization (TOCICO), a certification organization founded by Dr. Eli Goldratt. He later served as Director of Certification responsible for implementing TOCICO's certification program.

Now retired, **JOHN G. SCHLEIER**, Jr. was President and Chief Operating Officer of the Mortgage Services Division of Alltel, Inc., Executive Vice President of Computer Power, Inc., and Director of Office Systems and Data Delivery for IBM. In these positions, he directed major software development projects, sales administration, and financial functions. He was also Director of Information Systems for IBM's General Systems Division, where he provided oversight for Development Engineering, Manufacturing, and Headquarters systems. He developed information systems for manufacturing, sales, and IBM strategic planning functions and was winner of an IBM Outstanding Contribution Award. He was a regular lecturer on Strategic Planning at IBM Executive Briefing Centers over a period of 15 years, speaking to CEOs and top executives of major corporations. He frequently took consulting assignments

dealing with complex project management issues around the world.

He served on the faculty of The University of Georgia College of Business Administration as IBM Executive in Residence and later as Executive Professor of Management, serving on both the Management Information Systems and Production Operations Management faculties. Mr. Schleier holds TOCICO certification in all disciplines. He co-authored *Managing Operations: A Focus on Excellence*, a college text emphasizing TOC concepts (North River Press, 2003). He also published *Turkey Tales*, a children's book (Tate Publishing, 2010).

Selected Bibliography of Eliyahu M. Goldratt

James F. Cox III and John G. Schleier, Jr.

Entries are listed by type and year except where revision and newer editions are available. Several entries have been translated into a number of languages.

Books

Goldratt, E. M. and Cox, J. 1984. *The Goal: Excellence in Manufacturing*. Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. and Cox, J. 1986. *The Goal: A Process of Ongoing Improvement*. Rev. ed. Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. and Cox, J. 1992. *The Goal: A Process of Ongoing Improvement*. 2nd rev. ed. Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. and Cox, J. 2003. *The Goal: A Process of Ongoing Improvement*. 3rd ed. Great Barrington, MA: North River Press.

Goldratt, E. M. and Cox, J. 2000. *The Goal: A Process of Ongoing Improvement*. Audio book. Minneapolis, MN:

Highbridge Audio Book.

The Goal was made into two movies (the book story and a how-to version for training). These are listed under video movie/presentations.

Goldratt, E. M. and Fox, R. E. 1986. *The Race*. Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. 1990. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. 1991. *The Haystack Syndrome: Sifting Information Out of the Data Ocean*. Audio book. New Haven, CT: Avraham Y. Goldratt Institute.

Copyright © 2010 by James F. Cox III and John G. Schleier, Jr.

Goldratt, E. M. 1990. *What is this Thing called Theory of Constraints and How should it be Implemented?* Croton-on-Hudson, NY: North River Press.

Goldratt, E. M. 1992. *Late Night Discussions 1-12 with Alex and Jonah*. New Haven, CT: Avraham Y. Goldratt Institute.

Goldratt, E. M. 1994. *It's not Luck*. Great Barrington, MA: North River Press.

Goldratt, E. M. 1996. Production: *The TOC Way Work Book*. New Haven, CT: Avraham Y. Goldratt Institute.

Goldratt, E. M. 1997. *Critical Chain*. Great Barrington, MA: North River Press.

Goldratt, E. M. 1998. *Essays on the Theory of Constraints*. Great Barrington, MA: North River Press.

Goldratt, E. M. 1998. *Late Night Discussions on the Theory of Constraints*. Great Barrington, MA: North River Press.

Goldratt, E. M., Schragenheim, E. and Ptak, C. A. 2000. *Necessary but not Sufficient*. Great Barrington, MA: North River Press.

Goldratt, E. M. 2003. Production: *The TOC Way* including CD-ROM Simulator and Workbook. Revised edition. Great Barrington, MA: North River Press.

Goldratt, E. M. 2005. *Beyond the Goal: Eliyahu M. Goldratt Speaks on the Theory of Constraints. Your Coach in a Box*. New York: Gilden Audio. Audiobook: 8 CDs.

Goldratt, E. M. 2008. *The Choice*. Great Barrington, MA: North River Press.

Goldratt, E. M. 2009. *Isn't it Obvious?* Great Barrington, MA: North River Press.

Theory of Constraints Journal Articles

Goldratt, E. M. 1987. "[Chapter 1](#) hierarchical management—The inherent conflict," *The Theory of Constraints Journal* 1(1):1-17.

Goldratt, E. M. 1987. "A visit—Modine, the McHenry plant," *The Theory of Constraints Journal* 1(1):19-40.

- Goldratt, E. M. 1988. "[Chapter 2](#) laying the foundation," *The Theory of Constraints Journal* 1(2):1-20.
- Goldratt, E. M. 1988. "Apologia or in the move towards the third stage," *The Theory of Constraints Journal* 1(2):23-38.
- Goldratt, E. M. 1988. "[Chapter 3](#) the fundamental measurements," *The Theory of Constraints Journal* 1(3):1-21.
- Goldratt, E. M. 1988. "A visit—When quoted lead times are too long," *The Theory of Constraints Journal* 1(3):23-46.
- Goldratt, E. M. 1989. "[Chapter 4](#) the importance of a system's constraint," *The Theory of Constraints Journal* 1(4):1-12.
- Goldratt, E. M. 1989. "A visit—(fictional visit—real plants). Looking beyond the first stage: Just in Time," *The Theory of Constraints Journal* 1(4):13-46.
- Goldratt, E. M. 1989. "[Chapter 5](#) how complex are our systems?" *The Theory of Constraints Journal* 1(5):1-14.
- Goldratt, E. M. 1989. "Looking beyond the first stage—Just in Time: Part two," *The Theory of Constraints Journal* 1(5):15-48.
- Goldratt, E. M. 1990. "[Chapter 6](#) the paradigm shift," *The Theory of Constraints Journal* 1(6): 1-23.
- Goldratt, E. M. 1990. "Looking beyond the first stage—Just in Time: Part three," *The Theory of Constraints Journal* 1(6):25-43.

Journal/Magazine Articles

Goldratt, E. M. 1988. "Computerized shop floor scheduling," *International Journal of Production Research* 26(3):443-455.

Goldratt, E. M. 1993. "What is the Theory of Constraints?" *APICS—The Performance Advantage* June. Reprinted in *Selected Readings in Constraints Management*. Falls Church, VA: APICS. 1996, 3-6.

Goldratt, E. M. 1996. "My saga to improve production: Part 1," *APICS—The Performance Advantage* July 6(7):32-35.

And

Goldratt, E. M. 1996. "My saga to improve production: Part 2," *APICS—The Performance Advantage* August 6(8):34-36.

Reprinted in:

Goldratt, E. M. 1996. "My saga to improve production." Reprinted in *Selected Readings in Constraints Management*. Falls Church, VA: APICS 43-48. and

Goldratt, E. M. 2003. *Production: The TOC Way (Revised Edition) including CD-ROM simulator and workbook*. Great Barrington, MA: North River Press.

Goldratt, E. M. 1997. "The TOC approach to organizational empowerment," *APICS—The Performance Advantage* April 7(4):45-48.

Industry Week Late Night Discussion Series

Goldratt, E. M. 1991. "Late-night discussions I: Is your inventory putting you a continent away?" *Industry Week* July 1, 240(13):24-26.

Goldratt, E. M. 1991. "Late-night discussions II: Single-source purchasing's long-term effects can be devastating," *Industry Week* August 5, 240(15):29-31.

Goldratt, E. M. 1991. "Late-night discussions III: Transfer prices can be perilous, no matter how they're determined," *Industry Week* September 2, 240(17):68-70.

Goldratt, E. M. 1991. "Late-night discussions IV: Why lightless plants got buried under the carpet," *Industry Week* October 7, 240(19):55-57.

Goldratt, E. M. 1991. "Late-night discussions V: Searching for Japan's core statement: Manufacturing success of Japanese business," *Industry Week* November 4, 240(21):30-32.

Goldratt, E. M. 1991. "Late-night discussions VI: Time for Total Quality Management to confront the real issues," *Industry Week* December 2, 240(23):51-53.

Goldratt, E. M. 1992. "Late-night discussions VII: Why engineering is the key to competition," *Industry Week* January 6, 241(1):17-19.

- Goldratt, E. M. 1992. "Late-night discussions: VIII: When is a paradigm shift really a paradigm shift?" *Industry Week* February 3, 241(3):63-65.
- Goldratt, E. M. 1992. "Late-night discussions IX: Dealing with a market downturn," *Industry Week* March 2, 241(5):43-45.
- Goldratt, E. M. 1992. "Late-night discussions X: Different markets, different prices," *Industry Week* April 6, 241(7):58-60.
- Goldratt, E. M. 1992. "Late night discussions XI: Tearing down the walls of distrust," *Industry Week* May 4, 241(9):27-29.
- Goldratt, E. M. 1992. "Late-night discussions XII: How cost accounting can get in the way," *Industry Week* June 1, 241(11):38-40.
- Goldratt, E. M. 1996. "Empowerment: Misalignments between responsibility and authority," white paper. Accessed March 26, 2010 at <http://www.goldratt.com/empower.htm>
- Goldratt, E. M. 2000. "Project Management: The TOC Way, Tutor Guide and Workbook," including CD-ROM simulator. Unpublished. Roelofarendsveen, The Netherlands: A.Y.G.I. Ltd.
- Goldratt, E. M. 2009. "Standing on the Shoulders of Giants," *The Manufacturer*, June, accessed Feb. 4, 2010 at http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants

Management Skills Workshop Series (Workbooks)

Goldratt, E. M. 1995. *Management Skills Workshop: Sessions 1-5*. New Haven, CT: Avraham Y. Goldratt Institute.

Session 1: Resolving day-to-day conflicts.

Session 2: Dealing with half-baked solutions.

Session 3: Initiating skills: Addressing chronic conflicts.

Session 4: Delegation skills: Aligning authority with responsibility; Giving clear instructions.

Session 5: Team skills: Achieving ambitious targets.

Video Movie/Presentations

Goldratt, E. M. 1995. *The Goal*, Des Moines, IA: American Media Incorporated. Movie.

Goldratt, E. M. 1995. *The Goal: The How-To Version*, Des Moines, IA: American Media Incorporated. Movie.

Goldratt, E. M. 2000. *Deciding on TOC* (Video presentation DVD), Bedford, UK: Goldratt Marketing Group.

Goldratt Program Series (Video/DVD)

Goldratt, E. M. 1983. *The OPT CONCEPTS: Executive Video Course*. Milford, CT: Creative Output.

Learning Module 1. The OPT Way of Thinking

1. *The Goal of a Manufacturing Organization*
2. *The Unbalanced Plant*
3. *Bottleneck and Non-Bottleneck Resources*
4. *Basic Rules of Batch Sizing*

Learning Module 2: The Just-In-Time System and the OPT Rules

5. *Just-In-Time vs. Just-In-Case*
6. *OPT Rules Applied in Just-In-Time*
7. *The Path to Logical Ropes*

Learning Module 3: The Fallacy of Cost Accounting

8. *Performance Measurement, Part 1*
9. *Performance Measurement, Part 2*
10. *Determining Product Cost*
11. *Investment Justification*

Learning Module 4: The Logical Ropes of OPT

12. *Rules of Winning the Game*
13. *Identification of Bottlenecks*

14. *Master Schedule and Derived Schedule*

15. *Safeguarding the Schedule*

16. *OPT as a Productivity Tool*

Goldratt, E. M. 1983. *MRP vs. OPT - Software vs. Thoughtware - Program 1- What MRP Really Does*. Milford, CT: Creative Output.

Goldratt, E. M. 1983. *MRP vs. OPT - Software vs. Thoughtware - Program 2- Where MRP Goes Astray*. Milford, CT: Creative Output.

Goldratt, E. M. 1999. *Goldratt Satellite Program Sessions 1-8*. (Video series: 8 DVDs) Broadcast from Brummen, The Netherlands: Goldratt Satellite Program. Program Introduction.

Session 1: Operations.

Session 2: Finance & Measurements.

Session 3: Project Management and Engineering.

Session 4: Distribution and Supply Chain.

Session 5: Marketing.

Session 6: Achieving Buy-in and Sales.

Session 7: Managing People.

Session 8: Strategy & Tactics.

Self-Learning Computer Education Software Programs

Goldratt, E. M. 2001. *TOC Enterprise Wide: A Complete Self-Learning Program*. Bedford, UK: Goldratt Marketing Group. (Video series: 16 CD-ROMs).

Session 1: TOC on operations.
Session 2: TOC on finance and measurements.
Session 3: TOC on project management and engineering.
Session 4: TOC on distribution and supply chain.
Session 5: TOC on marketing.
Session 6: TOC on sales and buy-in.
Session 7: TOC on managing people.
Session 8: TOC on strategy and tactics.

Necessary and Sufficient Series

Goldratt, E. M. 2002. *Necessary and Sufficient Series Sessions 1-10*. Bedford, UK: Goldratt Marketing Group. (Video series: 10 CD-ROMs).

Session 1: The reasons for technology.
Session 2: The basic assumptions of TOC.
Session 3: A look into the rules of operations.
Session 4: A look into the rules of project management.
Session 5: A look into the rules of distribution.
Session 6: A look into measurements.
Session 7: The role of software.
Session 8: Implementing TOC as a holistic philosophy.
Session 9: Getting true consensus from top management.
Session 10: The offer: Clients, software providers and TOC.

TOC Insights Series. 4 Self-Learning Computer Software

Goldratt, E. M. and Goldratt, A. (R). 2003. *TOC Insights. 4 Self-Learning Computer Software*. Bedford, UK: Goldratt Marketing Group.

Insights into distribution and supply chain.

Insights into finance and measurements.

Insights into operations.

Insights into project management and engineering.

Chapters in Books

Goldratt, E. M. 1997. "Focusing on constraints, not costs." In *Rethinking the Future*, R. Gibson, ed. London: Nicholas Brealey Publishing Ltd.

Conference Proceedings/Video Proceedings/Presentations

Goldratt, E.M., 1980. Optimized production timetables: A revolutionary program for industry. In: *APICS 23rd Annual International Conference*, Falls Church, VA: APICS.

Goldratt, E.M., 1981. The unbalanced plant. In: *APICS 24th Annual International Conference Proceedings*, Falls Church, VA: APICS.

Goldratt, E. M. 1983. Cost accounting: The number one enemy of productivity. In: *APICS 26th International Conference Proceedings*. October. Falls Church, VA: APICS.

Goldratt, E. M. and Fox R. E. 1987. The Theory of Constraints. In: *APICS 30th Annual International Conference and Technical Exhibit*, St. Louis. Falls Church, VA: APICS.

Goldratt, E. M. 1996. Theory of Constraints in industry. (Keynote presentation) *APICS Constraints Management Symposium Proceedings*. April 17-19. Falls Church, VA: APICS.

Goldratt, E. M. 2000. Keynote: Necessary but not sufficient. APICS Constraints Management Technical Conference, Tampa, FL. Falls Church, VA: APICS.

Goldratt, E. M. 1998. On Saddam Hussein, milestones, and how the Theory of Constraints applies to project management. ManagementRoundtable.com

Goldratt, E. M. and G. Plossl. 1984. *A town without walls*. White paper, distributed during APICS 1984 International Conference in Las Vegas/USA

Keynote Presentations/Video Conference Presentation

Goldratt, E. M. 1997. JFL-1 The roots of TOC and the 3 cloud approach. In: *Video Conference Proceedings Jonah Upgrade Workshop*. New Haven, CT: Avraham Y. Goldratt Institute. Nov. 3-6.

Goldratt, E. M. 1997. JFL-16—Using the 3 cloud approach for buy-in. In: *Video Conference Proceedings Jonah*

Upgrade Workshop. New Haven, CT: Avraham Y. Goldratt Institute. Nov. 3-6.

Goldratt, E. M. 2001. Keynote: "Turning TOC into 'The thing to do,'" presentation at Founding TOCICO Founder Conference. Atlanta, GA: TOCICO, November 16-19.

Goldratt, E. M. 2003. Keynote: "Making TOC the main way: The Goldratt Group strategy & tactic tree and the viable vision process," presentation at 1st Annual Worldwide Gathering of TOC Professionals. Cambridge, UK: TOCICO, September 7-10.

Goldratt, E. M. 2004. Keynote: "What is different about TOC?" In: *Video Conference Proceedings of 2nd Annual Worldwide Gathering of TOC Professionals*. Miami, FL: TOCICO, October 23-26.

Goldratt, E. M. 2005. Keynote: "Success through simplicity," In: *Video Conference Proceedings of 3rd Annual Worldwide Gathering of TOC Professionals*. Barcelona, SP: TOCICO, November 13-16.

Goldratt, E. M. 2006. Keynote: "The economy of the world: Past and future," In: *Video Conference Proceedings of 4th Annual Worldwide Gathering of TOC Professionals*. Miami, FL: TOCICO November 4-7.

Goldratt, E. M. 2007. Keynote: "Freedom of choice," In: *Video Conference Proceedings of 5th Annual Worldwide Gathering of TOC Professionals*. Las Vegas: TOCICO, November 3-7.

Goldratt, E. M. 2008. Keynote: "What is TOC?" In: *Video Conference Proceedings of 6th Annual Worldwide*

Gathering of TOC Professionals. Las Vegas: TOCICO, November 1-4.

Goldratt, E. M. 2009. Keynote: "Standing on the shoulders of giants," In: *Video Conference Proceedings of North American Regional Conference.* Tacoma, WA: TOCICO, June 6-9.

Goldratt, E. M. 2009. Keynote: "Lessons learned: The power of cause-and-effect and TOC=focus," In: *Video Conference Proceedings of 7th Annual Worldwide Gathering of TOC Professionals.* Tokyo, JP: TOCICO, November 16-19.

The Goldratt Webcast Series

Goldratt, E. M. 2008. *The Goldratt Webcast Program on Project Management: Sessions 1-5.* (Video series: 5 sessions) Roelofarendsveen, The Netherlands: Goldratt Marketing Group.

Goldratt, E. M. 2008. *The Goldratt Webcast Program from make-to-stock to make-to-availability: Sessions 1-5.* (Video series: 5 sessions). Roelofarendsveen, The Netherlands: Goldratt Marketing Group.

Strategy and Tactic Trees

Goldratt, E. M. 2008. Pay-per-Click (PPC) S&T, Level 3, July.

Goldratt, E. M. 2008. Projects Company S&T, Level 5, July.

Goldratt, E. M. 2008. Retailer S&T, Level 5, July.

Goldratt, E. M. 2008. Consumer Goods Make-to-Stock (MTS) to Make-to-Availability (MTA) S&T, Level 5, September.

Goldratt, E. M. 2009. Manufacturing Make-to-Order (MTO) Reliable Rapid Response S&T, Level 5, May.

All the latest Goldratt approved S&Ts can be downloaded from Goldratt Research Labs with the free HARMONY S&T viewer from: <http://www.goldrattresearchlabs.com>

POOGI Forum Letter Series

Goldratt, E. M. 1999–2002. POOGI Forum Letters 1–14.
White papers on implementing TOC.

Letter 1: Moving the Organization - the TOC Way.

Letter 2: Enabling TOC to spread much faster.

Letter 3: The direction of the solution.

Letter 4: The solution.

Letter 5: The solution (continued).

Letter 6: When logic and emotion clash, which one wins?

Letter 7: Local implementation of a holistic approach is an oxymoron.

Letter 8: Presentation to top managers—Deciding on a holistic approach.

Letter 9: How to get results with TOC.

Letter 10: Necessary but not sufficient—[Chapter 12](#).

Letter 11: The story line of necessary but not sufficient.

Letter 12: How to implement a holistic approach bottom-up?—The problem.

Letter 13: How to move an organization bottom-up?—The solution.

Letter 14: What should be done to make TOC the main way for running organizations?

Accessed at: http://www.toc-goldratt.com/Theory_of_Constraints.php?cont=137

Plays

Goldratt, E. M. 1995. UnCommon Sense. The play (final revision), New Haven, CT: Avraham Y. Goldratt Institute. Shown in over two dozen different cities and sometimes part of “An Evening with Eli.”

Commercial Software

Goldratt, E. M. 1979. Optimized Production Technology (OPT®). Manufacturing scheduling software. Milford, CT: Creative Output.

Index

A

A-plants

Abbott, A.

ABC. *See* activity-based cost accounting

abnormal variation (red zone)

Abramov, E.

academics/researchers

accountability systems

accounting

- changing environment of

- lean

- for patients

- supply chain

- TOC literature lacking on

- TOC research needs in

Ackoff, Russell L.

acquisition decisions

action plan

action stage

Actively Synchronized Replenishment (ASR)

- alerts in

- benefits of

- buffer levels stratified in

- buffer profiles in

- business benefits of

- case study of

- compared to MRP

components of
decoupled explosion in
ERP system functionality and
geographically distributed items in
highly visible indicators in
implementation considerations in
lead time in
manufacturing environment considering
MRP attributes vs.
MRP compromises and
part and group trails
planning visibility of

activities

five types of
planned duration of
problem-solving
productivity vs.
profile
scheduling to time/completion of
upstream

activity time

determining
increasing planned
variability of

activity-based cost accounting (ABC)

activity-on-node project network

additional cause reservation

adjusting buffers

adoption barriers, TOC

Aggounne, R.

aggregate buffers

aggregate stock

aggregation, mathematics and

Aggregation Principle

aggression research

Ahearn, Mike

AIM. See Asian Institute of Management
Alain, S.
alerts, ASR
Almaguer, Zulema
Almqvist, R.
Ambitious Target Tree (*see also* prerequisite tree)
 as basic sequencing tool
 of Philippines
 of wedding plan
Amen, M.
American Production and Inventory Control Society (APICS)
analysis roadmap, for TOC
“and” connector
Andrews, C.
Ansoff, H. I
antisocial behavior research
APICS. See American Production and Inventory Control
Society
application solutions
APS system
Aquinas, Thomas
Argyris, M.
Ashkenas, R. N.
Asian Institute of Management (AIM)
ASR. See Actively Synchronized Replenishment
assembly lines
assignable cause
assumptions
 change challenges and
 change initiatives and
 of cloud
 Cloud method and
 conflict
 of CS
 defining
 EC/conflicts and

EC/dilemmas and
fire fighting cloud surfacing
fundamental
growth
Inner Dilemma Cloud and
necessary
necessary condition
parallel
patient
PRT/cloud and
strategic
strategies/tactics and
sufficiency
of TOC
TOC philosophical
attention span
Atwater, J. B.
Aubry, M.
auditing
 core conflicts in
 TOC used in
Austin, R. D.
Avot, I.

B

Backlund, A.
bad multitasking
Bailie, B.
Balakrishnan, J.
balance chart
balanced flow, TDD and
balanced scorecard (BSC)
Balderstone, S. J.

bank case

- PRT of
PRT/injection in
Snowflake approach

bank cloud

Barnard, A.

Barnes, R.

batches

- lead times and
peak/off-peak behaviors and

Baxendale, S. J.

Becker, C.

Becker, F.

Becker, S. W.

Beer, M.

behaviors

- antisocial research on
change/sustain desired
changing
chaotic project situations and
dysfunctional
human
peak/off-peak batch
sales
work

Behnke, L.

Belvedere, V.

Bennett, P.

Berry, R.

Best, W. D.

Betterton, C. E.

Bildson, R. A.

bill of material

- in MRP
aggregate

bio medical devices

black hole items
Blackstone, J.
bleak outlook, on CS
“The Blind Men and the Elephant”
blocking factors
Blue Ocean Strategy (Kim, C., Mauborgne)
BM. See Buffer Management
board game, supply chain
body of knowledge (BOK)
Bohr, Niels
BOK. See body of knowledge
Boorstin, D.
Bossidy, L.
bottlenecks
 efficiency increased in
 elevating permanent
 floating/multiple
 management of
 permanent
Boyd, John R.
Boyd, L.
Boyd, L. H.
Brailsford, S.
Branch diagram
breakeven chart
 of profit potential
 of volume exploitation
breakthrough injections
brewery, TOC’s 5FS applied at
“Broadening the Concept of Marketing”
Brocklesby, J.
Brooks, F. P.
Brown, D.
BSC. See balanced scorecard
budgets
 advantages/disadvantages of

buffers for
capital
date use for
flexible
master
process of
project
Throughput

buffer(s)

adjusting
aggregate
ASR stratifying levels of
for budgets
capacity control and
CCR and
controlling execution with
default
in distribution/replenishment solution
dynamic
equal time sections of
feeding
flow increase and
inventory and
inventory flow/buffer penetration and
personal schedule with
placement of
process flow and
production
profiles
project
resource
scheduling
smooth flow maintained and
space
state of the
target

target level
tasks with
for time management
of “to-do” list
tracking
variation

buffer burn rate

buffer consumption

continuous improvement with
flow disruptions influencing
rate of

buffer level profiling

Buffer Management (BM)

active

defining

doctor’s time scheduled and
establishing

flow management and
literature on
in MTA

ongoing improvement focus from
personal productivity increased with
process of
project control with
task prioritization in

buffer penetration

color codes used in
good enough levels using
inventory flow/buffers and
processing and

buffer size

initial set up of
production buffer and

buffer status

defining
orders released and

visible
buffered processes
buffered project plans
Build Innovation Empowerment Model
Built to Last (Collins, J.C.; Porras)
bullwhip effect
Burrell, G.
Burton-Houle, T.
business
 ASR benefits to
 CS impact on
 environment
 failure rate of
 for-profit owner perspective
 performance
 resource centric representation in
 restructuring
 theory of
business strategy
 compensation/reward policies in
 compromising factors in
 conflicting standards of performance
 defining
 emergent/deliberate
 5FS developing
 four strategy matrix for
 implementation process of
 improving
 marketing and
 performance standards conflicting in
 planning inadequate in
 of Porter, Michael
 resource-based view of
 sales and
 schools of thought on
 summary of schools

system analysis inability in
system conflicts disrupting
TA's role in
theories of
TOC for
business systems, with constraints
Buss, A. H.
Bust, Jeff
buy-in

Critical Chain management
layers of
marketing process steps for
motivation for
plus
process
sense of ownership in
steps of
TOC process of
up front

C

cadmium telluride (CdTe)
capacity
constraints
control
elevation
idle
load percentage and
management
planning
production
production orders and
protective

- seasonality and
- capacity buffers
 - defining
 - protective capacity and
- capacity constraint resource (CCR)
 - dealing with
 - scheduling/buffering of
- capital budgets

CARE. See Community Action for the Rehabilitation of Ex-Offenders Network

Carlson, B. J.

Cartier, B.

- case studies
 - of ASR
 - of constraint analysis workshop
 - of DBR
 - of VV

cash flow

categories of legitimate reservation (CLR)

categories of legitimate reservations (CLR)

causality clarity

causality existence

cause and effect

- diagram

- logic

- map

- necessary conditions and

- relationships

- strategy/tactic/parallel assumptions and

- terms/mapping protocol of

- tree

- UDE/idea relationship and

cause insufficient reservation

CC_G. See Critical Chain for Goods

CCMP. See Critical Chain Multiple Projects

CCPM. See Critical Chain Project Management

CCR. See capacity constraint resource

CCRT. See Communications Current Reality Tree

CC_s. See Critical Chain for Services

CdTe. See cadmium telluride

central warehouse (CWH)

certification, of leadership

chain analogy

Chakravorty, S.

challenges

- change

 - in Critical Chain

 - flow emphasis

 - improvement

 - of MRP

 - in project management

 - red curve

 - in service management

 - TOC

Chandrasekaren, S.

change

- agreement/to what to

- agreement/what to

- assumptions and

- becoming

- of behaviors

- challenging assumptions regarding

- consumption's sudden

- core conflicts of

- decisions of

- failure rate of

- fallacy of sudden

- good enough threshold and

- high failure rate of

- how to cause

- impact

implementing
inventory
Layers of Resistance to
low expectations for
managing
motives for
no urgency to
in personal productivity
policies/measurements of
POOGI measurement agreement achieving
purpose of
quantify impact of
questions about
results of
sequence
service management needs for
service organizations implementing
-sustain desired behavior
TA with
TOC analysis roadmap/proposed
what to
to what to

chaos theory

Charan, R.

Checkland, P.

cheetah items

Chen, J. C.

Cheng, C. H.

Chesapeake Consulting

The Choice (Goldratt, E. M.)

choopchicks

Christopher, Sharon Brown

chronic conflict

defining

resolving

chronic dilemmas

Church, A. H.
Church message
Churchill, Winston
Churchman, C. W.
Churchwell, L
CI. See continuous improvement
CIP. See continuous improvement process
clarity reservation
Clark, C. E.
classification scheme
client base, expanding
close call rate, response center
closed-loop MRP
Cloud (see *also* Evaporating Cloud)
 applications
 assumptions of
 bank
 conflict
 constructing valid
 core
 critical thinking skills from
 defining
 delay source solved with
 diagram
 dilemma
 flipping
 generic
 generic situations using
 in literature
 name calling
 necessary conditions utilized by
 needs/wants in
 operations conflict
 on playground
 problem-solving and
 PRT/assumptions and

win-win needed in
Cloud method (see *also* Evaporating Cloud)
assumptions/injections and
breakthrough solutions from
Consolidated Cloud as
Day-to-Day Conflict Cloud as
Fire-Fighting Cloud as
Inner Dilemma Cloud as
key points of
logical checks of
problem identified in
problem-solving with
solution construction in
solutions communicated in
storyline/building in
UDE Cloud as

CLR. See categories fo legitimate reservations; categories of
legitimate reservation

CM. See Cognitive Mapping

CMM. See Constraint Management Model

Cognitive Mapping (CM)

Cohen, Oded

collaborative execution

college student dilemma

Collins, J.

color codes

buffer penetration using
of work orders

Coman, A.

commission/omission, mistakes of
commitment

common cause

common costs

common denominator

communications

doctor/patient

EC dilemma of
notification system as
Communications Current Reality Tree (CCRT)
Community Action for the Rehabilitation of Ex-Offenders
Network (CARE)
comparative results, constraint analysis workshop
compensation
competitive advantage
 from Mafia Offer
 sustainable
Competitive Advantage (Porter, M. E.)
competitive edge
 building
 capitalizing on
 injections for
 Mafia Offers and
 premium
 reliability
 superior delivery for
completion times
complex environments
 guiding strategies for
 TA in
 TOC approach in
 tool selections in
complex organizations
 concepts in
 conflicts in
 core conflicts for
 flows in
 flow control
 measurements in
 problems in
 TDD in
 TOC/DBR/CCPM in
 UDEs of

- understanding
- complex systems
 - 5FS NS
 - physical constraint of solutions to
 - systems approach taken to UDEs in
- complexity
- compliance
- components
 - of ASR
 - lead times and MTA for
 - of ROI
- Conde, Ana Maria
- conflict cloud
- Conflict Resolution Diagram. See Evaporating Cloud
- conflicts
 - assumptions in
 - assumptions/EC and CCPM removing
 - CI
 - CI solutions for
 - in complex organization
 - day-to-day
 - in EC
 - EC/assumptions and
 - in father-son relationships
 - inherent simplicity and
 - Inner Dilemma Cloud assumptions causing
 - in metrics
 - with MRP
 - resolving chronic
 - between opposing actions
 - S&T identifying/removing
 - within system

systems vs. symptomatic
TOC's logistical solutions to
Connell, C.
consequences of actions
consolidated cloud
 core clouds relationship with
 multiple problems addressed with
 process of
 of production manager
consolidating process
constraint analysis workshop
 case study of
 current status of
 outcomes/comparative results from
 pilot project's current status of
Constraint Management Model (CMM)
 LTP's role in
 OODA Loop/TOC synthesis with
 seven step cyclical process of
constraints
 agreement on
 business systems with
 capacity
 of complex systems
 data accuracy at
 in discipline making
 elevating system
 health care system exploiting
 health care system identifying
 identifying
 internal operational
 in large-scale health care system
 markets
 non-
 as primary relevant factor
 scheduling

similar types of
subordinate to
TP/patient flow

Constraints Management Handbook (Cox, J. F., III; Spencer, M.)

consumer goods

consumption

of path slack

of project slack

sudden changes in

Conti, R. F.

continuous improvement (CI)

for conflicts

core conflicts in

DE's of

five questions for

focusing fear jeopardizing

Ford/Ohno lessons of

limiting/enabling paradigms in

management mistakes under pressure from

measurements/incentives of

TOC used in

continuous improvement process (CIP)

contractors/residents

control

mechanisms

processes

control points

product structure and

scheduling

convergence

CRT/inherent simplicity and

points

resource contention and

Conway, R.

co-op dilemma

Corbett, T.

CORE. See Cycle of Results

core cloud

core conflicts

- in auditing

- of change

- in CI

- for complex organizations

- defining

- EC and

- injections and

- injections breaking

- in organization

- service providers

- similar types of

- with solutions

- stakeholders identifying

- TP breaking through

core content, of TOC-Prisons

core problems

- identification of

- to positive actions

- UDE's in

Cormier, J. A.

Corpuz, Jenilyn

corrective actions

cost accounting

- activity-based

- business environment with

- development of

cost centers

cost control

cost world

costs, decision-relevant

cost-world paradigm

Cost-World Thinking

course materials, of TOC-Prisons

Covey, Stephen M. R.

Cox, J. F., III

CPM. See Critical Path Method

Critical Chain

benefits of

challenges implementing

CORE used in

defining

doctor's time scheduled and

failure causes of

flow control with

implementation empowerment model

implementation (step-by-step) of

implementation planning of

incomplete schedule/project buffer in

Lean working with

management buy-in

merging paths in

non-project work and

organization/purpose of

POOGI with

project management and

project network/time buffers and

project protection sources for

project schedule fully protected in

projects with

questions concerning

scheduling

in single project management

software's role in

task priorities in

three rules working together in

waiting and

Critical Chain (Goldratt, E. M.)

Critical Chain for Goods (CC_G)
Critical Chain for Services (CC_S)
Critical Chain Multiple Projects (CCMP)
Critical Chain Project Management (CCPM)
in complex organizations
conflict removed in
elements of
implementing
managerial actions/responsibilities supporting
project for ideas and
Critical Path Method (CPM)
multiple project management using
origins of
single project management using
single projects with
critical success factors (CSFs)
critical thinking skills
CRM. See customer relationship management
cross-functional team
CRT. See Current Reality Tree
CRT-EC-FRT method
CS. See customer support services
CSFs. See critical success factors
Csillag, J.
CTT. See customer tolerance time
culture
organizational
Throughput driven-
cumulative lead time
curfew dilemma
Current Reality Tree (CRT)
bank and
bottom up with
cause-effect relationships in
constructing

convergence and
CS with
developments in
health care facilities building
negative effects in
of Sheila's swimming
skeleton diagram of
with UDEs
warranty

customer label printer
customer relationship management (CRM)
customer support services (CS)
assumptions and
bleak outlook on
business impact of
change decisions and
CRT of
defining
differential pricing and
dilemma of
dilemma/assumptions of
expert service launching of
FSE visits of
as hostage
income erosion and
installations/implementations of
new environment for
problems facing
service offering of
TPM and
unnecessary services of
VAS of
warranty method recommendations of
warranty periods and
customer tolerance time (CTT)
customer UDE cloud

customers

- needs focused on
- predictable response to
- task flow toward
- value flowing toward
- value perception increased in

CWH. See central warehouse

cybernetic system

Cycle of Results (CORE)

- basic principles of
- Critical Chain using
- feedback loops in
- sales and
- Solution Selling and
- steps required for
- TOC practitioners group and

D

da Vinci, Leonardo

Daellenbach, H.

dampening, of stock buffers

Danos, G.

data accuracy, at constraints

Davies, J.

Davies, R.

Davis, K. R.

day-to-day conflicts

DBM. See Dynamic Buffer Management

DBR. See Drum-Buffer-Rope

DBR_G. See Drum-Buffer-Rope for Goods

DBR_S. See Drum-Buffer-Rope for Services

DCE. See decisive competitive edge

DDP. See due date performance

Decaluwe, L.
decision mechanism
decision steps, continued growth
decision-making process
 approaches to
 behavioral aspects of
 tools for
decision-relevant costs
decisive competitive edge (DCE)
Deckro, R. F.
decoupled explosion
Dedera, C. R.
dedicated durations
dedicated task times
default buffer
defects
defects per million opportunities (DPMO)
deficiencies, of organization
Define-Measure-Analyze-Design-Verify (DMADV)
Define-Measure-Analyze-Improve-Control (DMAIC)
delay source
demand
demand/supply
Demeulemeester, E.
Deming, W. E.
Demmy, B. S.
Demmy, W. S.
dental practice
dependency
 high degree of
 material/resource
 two types of
dependent resources
DEs. See desired effect
design decision, balance/unbalance
Design For Six Sigma (DFSS)

design process
desired effect (DEs)
detail complexity
Dettmer, H. W.
DFSS. See Design For Six Sigma
differential pricing
differentiation
dilemma cloud
dilemmas
 chronic
 college student
 co-op
 of CS
 CS/assumptions and
 curfew
 drinking age
 EC communications
 EC/assumptions
 in father-son relationships
 inner
 Las Vegas
 nurse's
 personal
 personal productivity
 poor grades
 rules
 training hours
 white-collar burnout
direct costing
direct costs
direct/variable costing
disaggregation
disciple making
Dismukes, J. P.
disruptions
disruptive students

distribution

- five questions applied to
- idea flows obligations in
- retail
- solution
- S&T
- TDD of
- TDD performance and
- TOC and
- workload of

divergence point

diversification

Divr, D.

DMADV. See Define-Measure-Analyze-Design-Verify

DMAIC. See Define-Measure-Analyze-Improve-Control

doctor unit time (DU)

doctor/patient communication

doctor's perspective

doctor's time scheduled

done, what is

double subordination

DPMO. See defects per million opportunities

drinking age dilemma

dropouts

Drucker, P. F.

drug trafficking

drum utilization

Drum-Buffer-Rope (DBR)

- 5FS used for

- background perspective on

- buffer of

- capacity constraints identified in

- case study of

- in complex organizations

- defining

- drum of

expected completion times in
flow emphasis challenging
flow management in
illustration of
in A-plants
in I-plants
in T-plants
in V-plants
OPT and
precursor literature of
problems with
PSTS with
rope in
scheduling literature on
simulations of
solution direction and
system of
traditional methodology and
Drum-Buffer-Rope for Goods (DBR_G)
Drum-Buffer-Rope for Services (DBR_S)
DU. See doctor unit time
Duclos, L. K.
due date performance (DDP)
due dates
Dumond, E. J.
Dumond, J.
Dunbar, R.
Duncker, Karl
Dweck, Carol
Dynamic Buffer Management (DBM)
 disabling
 SDCs using
 target levels increasing/decreasing in
 using
dynamic buffers

dysfunctional behaviors

E

Earl, Ezra

earned value reporting

Earned Value System (EVS)

EBQ. *See* Economic Batch Quantity

EC. *See* Evaporating Cloud

EC-CRT(B)-FRT(B)-NBR method

Economic Batch Quantity (EBQ)

Economic Order Quantity (EOQ)

ECPs. *See* engineering change proposals

Eden, C.

Edison, Thomas

educators

effective mechanism

80-20 rule

Einstein, Albert

Eisenstat, R. A.

elephant items

elevate step

Elton, M.

empowerment

end products

endangered need

engineering change proposals (ECPs)

engines of disharmony/harmony

Enterprise Project Management (EPM)

Enterprise Resources Planning (ERP)

ASR and

POOGI not supported by

Throughput-driven rules from

enterprise-level

entities

entity clarity

entity existence

environment

for CS

MTA's problematic

S-DBR fit and

S-DBR not suited for these

with S&T

EOQ. See Economic Order Quantity

EOQ(43) system

EPM. See Enterprise Project Management

ERP. See Enterprise Resources Planning

evaluation form

Evaporating Cloud (EC)

assumptions/conflicts and

assumptions/injections/core conflicts and

bank UDE and

college student dilemma and

communication dilemma in

conflict between opposing actions and

conflicts in

conflicts/assumptions and

constructing

defining

inherent conflict in

system performance and

template/hints for

training hours dilemma in

white-collar burnout dilemma and

EVS. See Earned Value System

execution

controlling

horizon

operations planning and

phase

S&T monitoring
Execution (Bossidy, Charan)
expectations
expected variation (green zone)
expert service launching
expertise
external reporting

F

Fabri, R.
failure
failure rate
family inclusion
Farah, K. S.
FASB. See Financial Accounting Standards Board
father-son relationships
 chronic dilemmas in
 conflicts in
 co-op dilemma in
 curfew dilemma in
 dilemmas in
 drinking age dilemma in
 Las Vegas dilemma in
 major issue dilemma in
 poor grades dilemma in
 rules dilemma in
Fawcett, S.
FCOs. See Field Change Orders
FDA. See Federal Drug Administration
Federal Drug Administration (FDA)
feedback
 accountability systems and
 government

information system and
loops
metrics and
of offenders
real-time system of
system
trainer

feeding buffers

Feldman, J. I.

fever charts

Field Change Orders (FCOs)

Field Service Engineer (FSE)

finance

TOC literature lacking on

TOC managing

TOC research needs in

Financial Accounting Standards Board (FASB)

Finch, B. J.

fire fighting

fire fighting cloud

assumptions surfaced in
building

logical connections in

problem identified in

solution constructed in

solutions communicated in

storyline/building in

First Solar Inc.

TOC contributions to

TOC's holistic implementation at

TP's role at

Five Focusing Steps (5FS)

Brewery using

business strategy developed through

complex systems and

DBR using

OODA Loop and
production operations with
of TOC
TOC strategy and
well-behaving organization from
5FS. See Five Focusing Steps
Flanders, Walter
flexible budgets
flipping cloud
floating bottlenecks
flow
 buffer increase maintaining
 cash
 centric representation
 in complex organization
 concentrating on
 control
 DBR challenge of
 disruptions
 focusing process on
 of ideas
 lines
 POOGI improving
 principle of
 of products
 reentrant
flow management
 in DBR
 execution/BM and
flying pig injection
focus
focusing matrix
focusing process
focusing step
Follet, Mary Parker
Ford, Henry

Ford production system
forecasts
 efficient
 misunderstandings of
 models impossible to find for
 rules of
Forgeson, S.
Foster, W. R.
Fox, R. E.
frame of reference, supply chain
Frazier, G.
free goods
Friend, J. K.
FRT. See Future Reality Tree
Fry, T. D.
FSE. See Field Service Engineer
FTPA. See Full Thinking Process Analysis
full kitting
full planned load
Full Thinking Process Analysis (FTPA)
Funcke-Bartz, Michael
functional management solutions
funnel experiment
funnel management
Future Reality Tree (FRT)
 bank
 injections in
 NBR and
 of Sheila's swimming
 S&T and

G

GAAP. See generally accepted accounting principles

gain sharing
Gandhi, Mahatma
Gantt, Henry L.
Gantt charts
Garcia, Marilyn
Gardiner, L.
Gardiner, S. C.
Gass, S.
Gauss, Carl Frederick
GDM. See Global Decision-Making
gedanken exercises
General Electric
General Motors
generally accepted accounting principles (GAAP)
generic cloud
generic strategy
generic structure, of S&T
geographically distributed items
Georgiadis, P.
getting started in TOC
Ghiselli, G.
Gibson, J.
Gillespie, J. R.
GKN Automotive
Glatter, Gila
Global Decision-Making (GDM)
global metrics
Gluxberg, Sam
The Goal (Goldratt, E. M., Cox, J.)
goal-orientated organization
goals

- detailed implementation plan for
- health care system defining
- knowledge base achieving
- long-term
- necessary conditions vs.

performance/gap between
of personal productivity
of projects
setting of

Goldratt, E. M.

The Choice by

Critical Chain by

The Haystack Syndrome by

It's Not Luck by

The Race by

satellite program of

"Standing on the Shoulders of Giants" by

success criteria recommended by

TA invented by

TP conceived by

Viable Vision of

VV goal originated by

Goldratt, Rami

good enough threshold

Good to Great (Collins, J.)

Goodrich, D. F.

Google Scholar

Gordon, T. M.

government

feedback

health care issues

perspective

Grando, A.

Granger, C. H.

Green, K. W.

Green, L.

green curve

Gregor, M.

Grinnell, John

Grosfeld-Nir, A.

growth

- assumptions
- curves/stability curves
- decision steps and injections
- growth matrix
- Grubb, Jeff
- Grubb, Orman
- Guan, Z. L.
- Guide, V. D.
- Gupta, M.
- Gupta, S. K.
- Gupta, Sanjeev

H

- Hamel, G.
- Hamilton, George “Chip”
- Hansen, Jesse
- harmony
- Harmony for creating S&T trees
- Harowitz, R.
- Harper, P.
- Harris, F. E.
- Harris, Jennifer
- Hart, Leslie
- Hasgul, S.
- The Haystack Syndrome* (Goldratt, E. M.)
- health care facilities
 - CRT building in
 - new core problem addressed in
 - process unit training of
 - sphere of influence in
- health care system
 - constraints exploited in

constraints identified in
goals defined of
governmental issues in
growth enhanced in
high level/lower value of
large-scale
Lean in
management priorities of
model of
improving patient flow through
POOGI started in
process flow of
process improvement in
socialized
TA in
TOC popular in
UDEs of
VV for

Healy, T. L.

heat treat operation

Herbein, W. C.

Herroelen, W.

high degree of dependency

Hilmola, O.-P.

Hinchman, J.

hit ratio

Hobbs, B

Hoel, K.

holistic distribution system

holistic implementation

decision mechanism for
of public sector TOC

S&T guiding

TOC

TOC 4x4

TOC/recommendations for

Hoover, Holly
hospital perspective
hostage, CS
Houle, D. T.
house on fire reservation
how to cause change
 PSTS and
 TOC - Prisons and
Howard, N.
HSM. *See Human Systems Management*
Huang, J.-Y.
Huang, S. H.
Huff, P.
Hughes, M. W.
human behavior
human element
Human Systems Management (HSM)
Hurley, S. F.
hybrid production method
hypothesis

I

I-plant research
I-plants
"I" to "They"
IDD. *See Inventory Dollar Days*
idea flows
idle capacity
IJPR. *See International Journal of Production Research*
image differentiation
Imai, Masaaki
Immelman, Ray
implementation

ASR considerations for
of change
of Critical Chain
Critical Chain planning for
CS with
details disagreement of
five step process of
goals and
of improvement program
of injections
injections/solution
input-output map of
MTA issues in
of POOGI
of practical segmentation
S-DBR issues of
TOC follow-through after
implementation process
of business strategy
TOC workshop for
improvement
challenges
challenges/limiting vs. enabling paradigms
gaps
implementing
large-scale health care system initiatives of
large-scale health care system needing
mistakes during
in personal productivity
potential
TOC tools of
incentives
income erosion
income statement
direct/variable costing on
traditional/Throughput

individual project system

industry solutions

inference

information system

inherent potential

inherent simplicity

Inherent Simplicity Ltd.

initiatives/projects

injections

bank PRT with

breakthrough

Cloud method and

for competitive edge

core conflicts and

core conflicts broken by

defining

EC and

in FRT

growth

implementing

to Inner Dilemma Cloud

IO map and

mini-project plan for

Negative Branch and

in personal dilemma

several/solution implementation and

Inman, R. A

inmates

Inner Dilemma Cloud

conflicting assumptions of

injections to

logical statement check of

sequence building

solution communicated in

solution creation of

storyline/building

inner dilemmas
input-output map
installations, CS
instructions, differentiated
insurance company perspective
integrated scheduling
interchangeable parts
Intermediate Objectives (IO) Map
internal operational constraints
internal reporting
International Journal of Production Research (IJPR)
International Transactions in Operational Research (ITOR)
inventory
 buffer placement of
 changes
 control
 determining
 distribution solution impact on
 excess
 flow/buffer penetration
 inventory value days and
 levels
 revenue vs.
 strategic positioning of
 turnover
 turns
Inventory Dollar Days (IDD)
inventory value days
investment centers
investments
InWEnt
IO map
Irlenusch, Bernd
ITOR. *See International Transactions in Operational Research*
It's Not Luck (Goldratt, E. M.)

J

Jackson, G. C.
Jackson, M. C.
James, G.
James, S. W.
Jamieson, N. R.
Jessop, W. N.
JIT. See just-in-time
Johanson, U.
Jones, Daniel
Jones, S.
JORS. See *Journal of the Operational Research Society*
Journal of the Operational Research Society (JORS)
just-in-time (JIT)

K

Kadipasaoglu, S.
Kadipasaoglu, S. N.
Kahn, K. B.
Kaizen Kaizen: The Key to Japan's Competitive Success
(Imai)
Kanban system
Kaplan, R. S.
Karan, K. R.
Kartal, K.
Kayton, D.
Kelley, J. E.
Kendall, Gerry
Kerr, Steven
Kerzner, H.

key performance indicators (KPIs)
Khumwala, B. M.
Kim, C.
Kim, S.
Kim, W. C.
King, R.
Klein, D.
Klusewitz, G.
Knight, A
knowledge base
knowledge organizer
Ko, H.-J.
Koljonen, E. L.
Koller, G.
Korte, G. J.
Kosturiak, J.
Kotler, P.
Kotter, J. P.
Koziol, D.
KPIs. See key performance indicators
Krishmaswamy, S.

L

labor costs
Lambrecht, M.
Lampel, J.
Land, M. J.
language limitations
large-scale health care system
 constraints in
 human element in
 improvement initiatives of
 improvements needed in

industry solutions adapted to
problems facing
problem-solving techniques in
safe platform/effective mechanism for
workforce of

Las Vegas dilemma

Lawrence, S. R.

Layers of Buy-In

Layers of Resistance

to change

no problems in

problem disagreement in

problem out of my control in

to solutions

Lea, B. R.

lead generation

lead time

in ASR

batches with

cumulative

elements of

end products/different

managed components with

manufacturing

realistic visibility of

transportation

leadership

certification

low-cost

spiritual

Leading Change (Kotter)

Lean

Critical Chain working with

defining

in health care system

methodologies integrated with

principles of
project environment disconnects with
project management
pursuit of perfection in
in system of systems
TOC accelerating
TOC implemented after
waste and
lean accounting
Lean principles
Lean Six Sigma (LSS)
design choices
project environment attitudes toward
TOC integration with
Lean Thinking (Womack, Jones, D.)
Lee, B.
Lenort, R.
Lepore, D.
Leshno, M.
LeTourneau Technologies, Inc
Lettiere, C. A.
Leus, R.
levels/steps relationship
Levison, W. A.
Levy, F. K.
Lewin, Kurt
life goals
achieving
thought processes achieving
TP attaining
Likert scale
Lindblom, C. E.
Lindsay, C. G.
literature. *See also* TOC literature
on BM
cloud

DBR
finance lacking
of project management
of service organizations
S&T vs. strategy in
on TP

load control

local efficiencies

local improvement/waste metrics

local measurements

local metrics

local operating expense metrics

logic, PA

logic branch

disruptive students and
as hopscotch

instructions differentiated with
in science

logic rules

logic template, S&T

logic tools

logical checks

logical connections

logical relationships

logical statements

Logical Thinking Process (LTP)

CMM with

long-term goals

Lorenz, Edward

Louw, L.

Low, J. T.

low-cost leadership

LSS. See Lean Six Sigma

LTP. See Logical Thinking Process

Luck, G.

Luebbe, R.

M

Mabin, V. J.

Mafia Offers

can you create
competitive edge and
creating
developing
preparing
problem agreement and
psychology of
sales increased by
solution agreement and
sustainable competitive advantage from
testing of
what it's not
who are recipients of

mainstream acceptance, TOC

mainstreaming

maintenance repair and overhaul (MRO)

major issue dilemma

make-or-buy

make-to-availability (MTA)

BM in

for components

generic issues in

implementation issues of

MTO mixed environment with

MTO/items fitting in

MTS relationship with

MTS/MTO moving to

problematic environments for

sales semi-continuous behavior with

software considerations of

make-to-order (MTO)

MTA/items fitting in
MTA/mixed environment with
MTA/moving from
sporadic demand managed as

make-to-stock (MTS)

MTA relationship with
MTA/moving from
special methodology required in
undesirable attributes of

Malcolm, D. G.

Malhotra, M. K.

management

active role needed by
of bottlenecks
capacity
chronic dilemma facing
of distribution/replenishment solution
funnel
health care system priorities of
material
mistakes
NBR and
omission/commission mistakes of
pipeline
principles
of product portfolios
resource
responsibilities
of SDCs
of stock
of TOC buy-in process
unstructured approaches of

Managing Operations: A Focus on Excellence (Cox;
Blackstone; Schleier)

manpower, underutilized

Mantel, S. J.

manufacturer/distributor

manufacturing

environments

environments/ASR consideration

lead time

priorities

reliability competitive edge in

S&T strategy of

Manufacturing at Warp Speed (Shragenheim/Dettmer)

manufacturing orders (MOs)

marketing

business strategy and

buy-in process with

defining

strategy

TOC managing

markets

constraints

development

expectations

penetration

segments

Mason, R. O.

Mason, Robert Award

master budgets

mastering the core

Matchar, D B.

material dependency

material flow

material management

material release

Material Requirements Planning (MRP)

ASR attributes vs.

ASR/compromises of

challenges of

closed-loop
compromises with
conflicts with
history of
organizational influence of
materials
acquisition decisions on
quantities held of
synchronization
mathematics, aggregation and
matrix structure
Matta, N. F.
Mauborgne, R.
M-B framework
McAdam, R.
McHugh, A.
McKay, K. N.
McMaster, Harold
McNamara, K.
mean, fallacy of
measurements
change and
of CI
in complex organizations
discarding local
of disciple making
local
of personal productivity
of projects
PSTS organizations requiring
purpose of
six general local
system
TDD usefulness of
Throughput using simple
TOC improvements of

Measuring and Managing Performance in Organizations
(Austin)

Mediate, B. A., Jr.

medical practice

Mentzer, J. T.

merging paths

meta-methodology, TOC

methodologies

- mapping

- TP using

metrics

- in conflict

- feedback/accountability system and

- local improvement/waste

- local operating expenses

- reliability

- speed/velocity

- stability

- strategic contribution

- TOC using

Meyer, Denise

Meyer, Theresa

Middleton, C. J.

military organizations

Miller, D. M.

Miller, J.

Miller, R. W.

Millstein, H. S.

Min, H.

Mingers, J

mini-project plan

Mintzberg, H.

mistakes

- of commission/omission

- improvements with

- management

Mitroff, I. I.
money making box
Moore, R.
Morgan, G.
Morin, C.
Morris, J. S.
Morris, R. C.
Morton, T. E.
MOs. See manufacturing orders
Moseley, S. A.
Moss, H. K.
motion, excess
motivation, for buy-in
Motwani, J.
MRO. See maintenance repair and overhaul
MRP. See Material Requirements Planning
MRP II systems
MTA. See make-to-availability
MTO. See make-to-order
MTS. See make-to-stock
multiple bottlenecks
multiple project environments
 bad multi-tasking in
 four systems of
 scheduling projects in
multiple project Gedankens
multiple project management
multitasking
 bad
 waiting during
Munro, I.
Murakami, S.
Muris, Fiet
Murphy, R.

N

National Trade Union Congress (NTUC)

NBR. See Negative Branch

necessary assumptions

necessary conditions

necessity

necessity logic

needs/wants

- alternatives to

- customers/focusing on

- differentiating between

- identifying underlying

- validating

Negative Branch

- defining

- diagram

- handling process of

- injections and

- negative outcomes trimmed in

- obstacle difference with

- as predictive tool

- solution structure of

Negative Branch Reservation (NBR)

- daily problem-solving with

- FRT and

- managers/co-workers and

- using

negative effects

negative peer pressure

Neimat, T.

nervousness

net present value (NPV)

new core problems

- new solution
- Newton, Sir Isaac
- Ning, J. H.
- Nolan, Jim
- non-constraints
- non-contractual performance reports
- non-critical path
- non-project work
- nonstandard application, of TOC
- No-Questions-Asked policy (NQA)
- normal variation (yellow zone)
- Norris/AOT
- Norton, D. P.
- notification system
- NPV. See net present value
- NQA. See No-Questions-Asked policy
- NTUC. See National Trade Union Congress
- nurse's dilemma

O

- observation
- observe/orient/decide/act. See OODA Loop
- obstacles
 - addressing
 - Negative Branch difference with
 - PRT identifying
- Odom, R.
- offenders
 - evaluation form
 - feedback of
 - Likert scale
 - negative peer pressure obstacle to
 - work important to

Ohno, Taiichi

ongoing improvement

BM focus on

fundamental questions for

S&T and

OODA Loop

CMM/TOC synthesis with

fast cycles in

5FS steps and

operating expense

operating profits

operational improvements

operational system

operations conflict cloud

operations planning

opportunities

limiting

sales and

wasting

OPT. See Optimized Production Technology

Optimized Production Technology (OPT)

order lead time

order point system

order priority status

order release

order spike protection, qualified

Oregon Freeze Dry

organizations

chain

core conflict in

with Critical Chain

deficiencies of

existing systems/measures removed from

5FS improving

four levels of

goal-orientated

goals/S&T and
improvement gaps and
internal TOC champion required in
public sector vs. private sector
systems approach not adopted by
systems approach to
orientation step
Orlicky, J.
OR/MS structured approaches
outsourcing proposals
overhead costs
overproduction

P

Page, D. C.
Paige, H. W.
paradigms
 in CI
 cost-world
 limiting vs. enabling
 shift in
 throughput-world
parallel assumptions
Pareto's law
Park, Y. H.
Parkinson, C. Northcote
Parkinson's Law
part traits
parts shortages
Pass, S.
path slack, consumption of
patients
 accounting for

- assumptions and care
- due date setting of flow/constraints/TP and perspective
- selling services to transport scheduling
- value stream map and

Patterson, J. W.

Patwardhan, M. B.

pay per click

PDCA. See Plan-Do-Check-Act

peak/off-peak behaviors

Peng, Y. F.

Penvoisé, P.

People with Disabilities (PWDs)

performance

- auditing

- business

- business strategy and conflicting standards of criteria of

- distributors

- evaluation

- gaps/variation in

- global measures of

- goals/gap between

- measurement system

- measures

- scorecard, balanced

- standards

- superior

- value metric tracking

periodic reporting

permanent bottlenecks

personal dilemma

personal productivity

- BM increasing

- change in

- dilemma

- goals/strategies/measures of

- improving on

- time management and

- tools improving

PERT. See Program Evaluation and Review Technique

Peterson, J.

Petrini, A. B.

PFDs. See product flow diagrams

Phil, Greg

Philipoom, P. R.

philosophical assumptions, TOC

philosophical basis, TP

pilot projects

- constraint analysis workshop current status

- of distribution/replenishment solution

pilot study, TOC-Prisons

PIMS. See Profit Impact of Market Strategies

Pinedo, M

Pink, D.

Pinto, J. K.

pipeline control

pipeline management

pipelining

Pirasteh, R. M.

Pitagorsky, G.

Pittman, P. H.

Plan-Do-Check-Act (PDCA)

planned activity duration

planned load

- full

- short-term

planning

ASR visibility of
cycle
inadequate
problems process of
rules of
S-DBR procedure of
short-term
plant warehouse (PWH)
playground
Pliskin, J. S.
PLM. See Product Lifecycle Management
plus buy-in
PMBOK. See project management body of knowledge
PMO. See Project Management Office
pockets of excellence
Pocock, J. W.
Politou, A.
POOGI. See Process of Ongoing Improvement
poor grades dilemma
Porras, J. I.
Porter, M. E.
Porter, Michael
portfolio of projects system
POs. See purchase orders
positive actions
potato experiment
PP&E. See Property, Plant and Equipment
practical segmentation
Prahalad, C. K.
predictability
predicted effect existence reservation
predicted effects
predicted undesirable effects (PUDEs)
predictive tool
premium competitive edge
premium offer design

premium sales

Prerequisite Tree (PRT)

of bank case

cloud assumptions and

injections in

IO map/obstacles and

obstacles identified by

of Sheila's swimming

S&T and

Prescott, D. P.

presentation design

price-quantity curve

pricing indifference model

priorities changing

priority control

priority planning

prison officers

private sector

proactive, reactive vs.

problems

Cloud method identifying

in complex organizations

consolidated cloud addressing multiple

daily/U-shape solving

with DBR

facing CS

facing educators

fire fighting cloud identifying

investigation of

large-scale health care system facing

Layers of Resistance and

Mafia Offer/agreeing on

planning process for

resistance to

with SCM

scrap

to solutions implementation
solutions solving
U-shape solving
problem-solving
activities
approaches to
cloud applications for
with Cloud method
large-scale health care techniques in
methods/activity in
NBR used for
systems approach to
with TOC
tools for
TP's relationship to
U-shape for
Problem-Structuring Methods (PSMs)
process flow
buffer recovery and
of health care system
process improvement
process management, sequenced
Process of Ongoing Improvement (POOGI)
change measurement agreement achieving
with Critical Chain
ERP not supporting
flow improvement and
functional management solutions and
health care organization starting on
implementing
improvement over time with
sales with
processes
buffer penetration and
non-value added
projects vs.

ProChain
product

- development
- differentiation
- mix
- portfolio managing of
- supply chain for

product flow

- diagram
- in V plants
- resources and

product flow diagrams (PFDs)

Product Lifecycle Management (PLM)

product structure

- control points and
- resource information and

production

- activity vs.
- buffer
- capacity
- environments
- floor scheduling of
- lead time

production manager

production operations

- control needed in
- 5FS and
- time buffers needed in
- variability in

production orders

profession, in TOC

Professional, Scientific, and Technical Services (PSTS)

- DBR applied to
- expertise/assets of
- measurements required in
- service delivery of

strategies of
TOC challenges in
profit
center
potential/breakeven chart
TOC maximizing
Profit Impact of Market Strategies (PIMS)
profitability
Program Evaluation and Review Technique (PERT)
multiple project management using
origins of
single project management using
single projects with
project(s)
budgeting
buffers
business performance links with
companies
control
with Critical Chain
goals/objectives/measures of
plans
priorities
processes vs.
protection
reporting
resource contention and
resource priorities across
scope definition of
situations
slack/early consumption
tasks
tasks/lead times
time-traps in
project environment
Lean disconnects with

LSS attitudes in
system improvements in
system of systems in
systems aligned in
waste in

Project Leadership Model

project management

challenges in

control mechanisms in

Critical Chain and

defining

dilemma cloud of

execution

failures in

five-step approach to

guideline development in

lean/traditional

literature of

multiple project management

pipe lining

plan objective issues in

single project management

sustaining Critical Chain

tactics/actions considered by

project management body of knowledge (PMBOK)

Project Management Institute

Project Management Office (PMO)

project network

activity-on-node

developing

resource contention and

project schedule

fully protected

in multiple project environments

resource-leveled

Projects S&T

partial structure/four levels and
processes essential for
WIP reduced from
Property, Plant and Equipment (PP&E)
protective capacity
PRT. See Prerequisite Tree
PSMs. See Problem-Structuring Methods
PSTS. See Professional, Scientific, and Technical Services
psychological barriers, to solutions
psychology, of Mafia Offer
public sector
 complicating factors in
 future TOC applications in
 S&T (created in harmony) in
 TOC holistic implementation in
PUDEs. See predicted undesirable effects
pull replenishment system
pull supply chain
pull system
pull-based demand generation
purchase orders (POs)
purchasing decisions
 acquisition decisions and
 outsourcing proposals and
pursuit of perfection
push, push, push syndrome
push system
PWDs. See People with Disabilities
PWH. See plant warehouse

Q

QFD. See quality function deployment
quality

improvements
service organizations enhancements of
quality function deployment (QFD)
Quinn, J. B.

R

RACE. See return on average capital employed
The Race (Goldratt, E. M., Fox, R. E.)
Radovilsky, Z. D.
Rahman, S-U
Rand, G. K.
Rapid Response
Rashidi, Hajah Ahmad
rate-based exploitation
Rational Analysis for a Problematic World (Rosenhead)
raw materials
Raz, T.
Razzak, M. A.
Reaching the Goal (Ricketts)
reactive, proactive vs.
reading test scores
realistic lead time visibility
reality, desired future to
real-time feedback system
reason code analysis
recoverable manufacturing
red curve challenge
reentrant flows
regional warehouse (RWH)
Reid, R. A.
Reimer, G.
relay runner
reliability competitive edge

reliability metrics
reliable rapid response
reliable replenishment time
reorder point systems
replenishment
 frequency of
 lead time (RLT)
 orders
 system
Replenishment fo Goods (R_G)
Replenishment for Services (R_S)
requests for proposals (RFPs)
Rerick, R.
research
 aggression
 antisocial behavior
 A-plant
 I-plant
 limitations of
 V-plant
 TOC accounting needs of
 TOCfE ongoing
reservations, overcoming
resource(s)
 activity profile of
 allocation
 -based business strategy
 buffers
 capabilities added and
 centric viewpoint
 dependency
 information
 leveled project schedule
 leveling
 management

management level
notification
priorities
product flows through
production capacity of
unbalanced capacities of
underutilized
unsynchronized
variability
WIP profile of
resource constraint
rate-based exploitation of
scheduling of
resource contention
activity-on-node project network and
common resource variability and
priority planning and
projects and
resolving
variability and
variability/convergence and
respect
response center close call rate
response time reductions
responsiveness
retailers
S&T for
S&T level 2 for
S&T level 3 for
S&T level 4-5 for
return on average capital employed (RACE)
return on investment (ROI)
classification clashes and
components of
revenue
reward system

Reyes, Miquel Perez
Reyes, P.
RFPs. See requests for proposals
 R_G . See Replenishment fo Goods
Ricardo, David
Ricketts, J. A.
Riezebos, J.
Rippenhagen, C.
risks, of solutions
Ritson, N.
Rizzo, T.
RLT. See replenishment lead time
roadrunner work ethic
Robinson, D. E.
Roby, Doug
ROI. See return on investment
Ronen, B.
root causes
Rose, E.
Roseboom, J. H.
Rosenhead, J.
Roybal, H.
 R_S . See Replenishment for Services
rules dilemma
Rules of Reasoning in Philosophy
Russell, G. R.
RWH. See regional warehouse

S

safe dates
 determining
 sales quotes earlier than
 special orders and

safe platform

Sale, M.

Sale, M. L.

sales

abolish local efficiencies

business strategy and

CORE and

cycle in days

disruptions to

execution

improving flow

Mafia Offer increasing

manager

market constraint

MTA/semi-continuous behavior of

operations conflict cloud vs

opportunities insufficient and

with POOGI

premium

project buffers and

psychology of mafia offer

quotes

SCM and

support

target levels and

templates

TOC managing

work flow of

sales funnel

introducing requests to

opportunities limited vs.

Samolejova, A.

sandbagging

Santiago, Cora

Sarbanes-Oxley Act of 2002 (SOX)

Sarria-Santamera, A.

satellite program, of Goldratt
Saxe, John Godfrey
SCA. See Strategic Choice Approach
Schaefers, J.
schedule reserves
scheduling

- buffers
- of CCR
- constraints
- control point
- Critical Chain
- DBR literature on
- discarding local
- integrated
- in multiple project environments
- non-constraints
- patient transport
- of production floor
- of resource constraint
- resources
- in single project management
- to time

Scheinkopf, L.
Schleier, J. G.
Schoemaker, T. E.
Schol, John
Scholes, J.
Scholl, A.
Schön, D.
Schonberger, R. J.
schools of thought
Schrageheim, E.
Schultz, Howard
Schultz, Kenneth "Ken"
Schwartz, C.
science, logic branch

scientific method

SCM. See Supply Chain Management

SCORE. See Singapore Corporation of Rehabilitative Enterprises

scrap problem

S-DBR. See Simplified Drum-Buffer-Rope

SDCs. See Sharp Demand Changes

seasonality

- capacity and

- distribution/replenishment solution managing

- stock management and

self-aware systems

self-regulation

selling prices

Senge, P. M.

sense of ownership, in buy-in

sensitivity analysis

sequenced process management

sequence-dependent-setup

sequencing tool

service environment

Service Level Agreements (SLAs)

service management

- challenges in

- change needs of

- unique characteristics of

service offerings, CS

service organizations

- change implemented in

- customer support services

- double subordination instilled in

- performance measures of

- professional/scientific/technical

- How to Cause the Change

- What to Change

- What to Change to

quality enhancements of
response time reductions in
TOC literature of
TOC popularity with
TOC steps for
value enhancement of
service providers
Service Science, Management, Engineering, and Design
(SSMED)
services
 differentiation
 turnaround time of
Sha, D. Y.
Shao, X. Y.
Sharp Demand Changes (SDCs)
 adjustments to
 DBM used with
 handling of
 management steps of
 significance of
Shaw, D.
Shewhart, Walter
Shi, J.
Shingo, Shingeo
shipping buffer
shorter-delivery orders
short-term objectives
short-term planned load
show stoppers
silver bullets
Simatupang, T. M.
Simons, J. V., Jr.
simple measures
Simplified Drum-Buffer-Rope (S-DBR)
 environmental fit of
 environments not suited for

implementation issues/processes of
planning procedure in
Simpson, W. P.
Simpson, W. P., III
Sims, D.
simulations
 of DBR
 of distribution/replenishment solutions
 drawbacks of
 TOC
Sinacka-Kubik, Edyta
Singapore Corporation of Rehabilitative Enterprises (SCORE)
single project Gedankens
single project management
 CPM/PERT critical paths in
 Critical Chain in
 scheduling in
single projects
Sirias, Danilo
Six Sigma methodology
 defining
 methodologies integrated with
 TOC accelerating
 TOC implemented after
skewed distribution
Skoog, M.
SKU. See stock keeping unit
slack
SLAs. See Service Level Agreements
Slevin, D. P.
Small, Belinda
Smith, G. R.
Smith, L. B.
Smith, Manfred
“Snowflake method”
social barriers, to solutions

socialized health care

Soft OR methods

Soft Systems Methodology (SSM)

software

Critical Chain and

distribution/replenishment solution assistance of

MTA/considerations of

Sohn, Bruce

solid waste management

buildup and

chain/system

UDE's in

Solution Selling

solutions

can't implement

Cloud method communicating

Cloud method with

to complex systems

core conflicts with

DBR and

details disagreements of

development of

direction agreements of

disagreement on

fire fighting cloud communicating

fire fighting cloud constructing

ingredients of

injections/implementation of

Inner Dilemma Cloud communicating

Inner Dilemma Cloud with

Mafia Offer agreeing on

Negative Branch structure of

negative ramifications of

problems solved by

risks of

social/psychological barriers to

from S&T
tools for
UDE cloud constructing/communicating

Sonawane, R.
SOX. See Sarbanes-Oxley Act of 2002
space buffers
Spangler, Todd
Spearman, M L.
special methodology
special or assignable cause
special orders
specialized applications, TOCfE
specific throughput
Spector, Y.
speed metrics
The Speed of Trust (Covey)
Spencer, M. S.
sphere of influence
spiritual leadership
sporadic demand
Sridharan, R.
Srikanth, M. L.
Srinivasan, M. M.
SSM. See Soft Systems Methodology
SSMED. See Service Science, Management, Engineering,
and Design
S&T. See Strategy and Tactic tree
stability cloud
 growth assumptions vs.
 growth injections (potential) vs.
stability curves
stability metrics
stakeholders
 core conflict identified by
 strategy sessions feedback of
“Standing on the Shoulders of Giants”

star items
starvation
state of the buffer
statistical fluctuation
statistics
status zones
Steele, D. C.
steeling
Stein, R. E.
Stephens, A. A.
steps/levels relationship
Steyn, H.
stigmatization
stock
 available/on-hand
 builddown of
 buildup of
 management confusion of
 management/seasonality and
 outs
stock buffers
 dampening influence on
 demand/supply/replenishment lead time and
 manufacturing priorities and
 structure of
 supply chain with
 target levels and
 zones
stock keeping unit (SKU)
Stoltman, J. J.
Stone, Tom
strategic assumptions
Strategic Choice Approach (SCA)
strategic constraints
strategic contribution metrics
strategic gating

strategic inventory positioning
strategic segmentation
strategy

assumptions with
assumptions/tactics
cause-effect and
common denominator defining
criteria for
for complex environments
deployment
desirable influence from
human behavior influencing
as journey
matrix
of personal productivity
OODA loop
planning cycle and
prerequisite conditions achieving
of PSTS
role of Throughput Accounting in
sessions/stakeholder feedback
TOC and the OODA loop

Strategy and Tactic tree (S&T)

application of
base growth from
basic structure of
benefits of
communication/synchronization with
concepts in creation of
conflict cloud and
conflicts identified/removed using
consumer goods
CRT/FRT/PRT/TRT and
defining
different environments with
distribution

distribution solution
elements of
execution monitored using
execution of
first step/goal of
four generic
generic structure of
harmony achieved using
holistic implementation guided by
implementing
layers of detail in
logic template for
lower levels of
manufacturing strategy using
NA connections of
ongoing improvement process and
organizational goals achieved through
pay per click
project companies using
public sector harmonious
reliable rapid response
retail
for retail
retail level 2 of
retail level 3 of
retail level 4-5 of
solutions from
strategy description for
strategy literature compared to
structure content of
structure details of
templates created for
three levels of
top of
TP analysis to
TPs cross reference with

TPs implementing
using TPs to implement
VV achieved through
stratification, of time buffer
structure details, S&T
student syndrome
subordination
success criteria
sucker rods
Sudden Demand increase/decrease
sufficiency
 assumption
 -base logic
 of breakthrough injections
Sufficiency Assumption
sufficiency gaps
Sullivan, T. T.
Sun Tzu
superior delivery
superior performance
suppliers practice
supply chain
 accounting
 board game
 concepts of
 for finished product
 five questions applied to
 frame of reference of
 pull
 push vs. pull of
 stock buffers across
 typical
Supply Chain Management (SCM)
 current problems with
 identifying problems in
 sales and

Supply Chain Management at Warp Speed (Schragenheim; Dettmer; Patterson)

supply factor

surrounding processes

sustainability

sustainable competitive advantage

synchronization

Synchronized Supply Chain Application

Synchronous Manufacturing (Umble, M.; Srikanth)

syntax guidelines, logical statements

system constraints

agreement on

exploiting

identifying

systems

analysis

application solutions of

basic diagram of

characteristics of

concept

defining

improvements/in project environment

levels

levels/vertical hierarchy

performance of

project environment with aligned

replenishment vs.

symptomatic conflicts

of systems

of systems/Lean in

whole view of

systems approach

to complex systems

five focusing steps in

to organizations

organizations not adopting

to problem-solving
TOC
TOC tools used in

T

T-plants

TA. See Throughput Accounting

tactical gating mechanism

tactics

TA_G. See Throughput Accounting for Goods

takt time

Taormina, Sheila

target level

- buffers/stock buffers

- DBM/increasing/decreasing

- sales and

target market

TA_S. See Throughput Accounting for Services

task management

- establishing

- system

task priorities

- in buffer management

- in Critical Chain

tasks

- buffering within

- duration

- duration estimates

- flow/toward customer

- skewed distribution and

- time and

Taylor, Audrey

Taylor, Frederick

Taylor, L. J., III
Taylor, L. T., III
Taylor, S. G.
Tayner, T.
TCO. See total cost of ownership
TDD. See Throughput Dollar Days
teaching techniques
Teaño, Adora
templates
Teyner, T.
Theory for Inventive Problem Solving (TRIZ)
theory of business
Theory of Constraints (TOC)
 academics/researchers role in
 accounting/finance research needs
 adoption barriers of
 analysis roadmap of
 analysis roadmap/proposed changes and
 background of
 background/holistic implementation of
 basic assumptions of
 benefits of
 bottom-up implementation of
 business strategy through
 chain analogy in
 CI using
 CMM/OODA Loop synthesis with
 in complex environments
 conflict logistics solutions of
 contributions to
 distribution and
 distribution/replenishment solution in
 finance managed with
 First Solar's benefits from
 First Solar's holistic implementation of
 five focusing steps (5FS) of

five questions of
for education
future public sector applications of
future research of
gaps/complexities in
healthcare organizations and
holistic implementation recommendations of
implementation process workshop and
implementation/follow-through of
improvement tools of
in prisons
in professional/scientific/technical services
Lean/Six Sigma accelerated by
Lean/Six Sigma methodology and
logic tools used in
LSS design choices with
LSS integration with
mainstream acceptance of
management principles/applications of
marketing/sales managed by
measures improvements from
meta-methodology of
metrics used in
nonstandard application of
organizational internal experts of
philosophical assumptions in
planning/control/sensitivity analysis in
practitioners getting started in
precursor literature of
private sector holistic implementation of
problem-solving with
profession in
profit maximizing in
PSTS challenges for
public sector holistic implementation of
recommendations for

results with
results/success
self-analysis of
service organizations popularity of
service organizations steps for
simulation games and
S&T TP and
subject matter expert of
systems approach paradigm shift with
systems approach/tools of
three basic elements of
Throughput focus of
TOCLSS (fully integrated)
train-the-trainer in
understanding lacking of
U-shape/assumptions of
vignettes

Theory of Constraints International Certification
Organization (TOCICO)

Thinking Across the Curriculum

Thinking Processes (TP)

CIP using
classificatory mapping
core conflicts resolved by
defining
development/use of
First Solar's use of
Goldratt conceiving of
history of
integrated
life goals attained with
literature on
methodologies used in
nature/use revisited of
patient flow constraints identified by
philosophical basis of

problem-solving activities relationship to
S&T and
S&T cross reference with
S&T implemented with
tool orientation
tool usage of
tools/purposes/relationships of
UDEs surfaced by
using

third-party maintenance (TPM)

Thompson, G. L.

Three-Cloud

analysis
approach
method

Throughput

budget prepared for
defining
-driven culture
-driven rules
First Solar's annual
holistic distribution system increasing
impact levels on
income statement
market segments and
per order
simple measures used for
specific
TDD and
TOC focus of
value days
-world paradigm

Throughput Accounting (TA)

for all methods
business strategy's role of
in complex environments

Goldratt inventing
health care system and
percent changes and
performance evaluation and
Throughput Accounting for Goods (TA_G)
Throughput Accounting for Services (TA_S)
Throughput Dollar Days (TDD)
alternatives to
balanced flow from
in complex organizations
distributors performance of
measurement usefulness of
Throughput and
Throughput per shelf space (TPS)
Throughput world, cost world vs.
time buffers
Critical Chain project network with
production operations needing
size of
stratification of
time management
buffers for
personal productivity and
time to reliably replenish (TRR)
time-traps
TMG. See Too Much Green
TMR. See Too Much Red
to what to change
TOC. See Theory of Constraints
TOC buy-in process
TOC distribution paradigm
TOC distribution/replenishment solutions
buffers in
defining
managing

pilot project of
results of
seasonality managed of
simulation of
software assistance in
testing of

TOC for education (TOC_{fE})

Ambitious Target Tree used in
cloud used in
logic branch used in
ongoing improvements/research and
reading test scores and
specialized applications of

TOC for Goods (TOC_G)

TOC for Services (TOC_S)

TOC implementation

4x4 holistic launch of
X-Y syndrome of

TOC Information Systems (TOCIS)

TOC literature

accounting/finance lack of
issues emerging from
nature of
of service organizations
special cases in

TOC mindset management workshop

TOC practitioners group

TOC Priority Management

TOC strategy

defining
different formats of
distribution/retail and
5FS and
foundations of
goals/conditions of

TOC_{fE}. See TOC for education

TOC_G. See TOC for Goods

TOCICO. See Theory of Constraints International Certification Organization

TOCIS. See TOC Information Systems

TOC-Prisons

- core content of
- course materials of
- family inclusion in
- future recommendations and
- at home
- necessity logic taught in
- negative peer pressure and
- pilot study results of
- reasons for doing
- respect and
- results of
- stigmatization and
- sufficiency logic taught in
- teaching techniques
- to what to change
- what to change in
- at work

TOC_s. See TOC for Services

“Today and Tomorrow”

“to-do” list buffer

Too Much Green (TMG)

Too Much Red (TMR)

tools

- selection of
- of TP
- for variability

TOs. See transfer orders

total cost of ownership (TCO)

Total Quality Management (TQM)

totally variable cost (TVC)
Toyota, Eliji
Toyota Production System (TPS)
“Toyota Production System Beyond Large-Scale Production”
TP. See Thinking Process
TPM. See third-party maintenance
TPS. See Throughput per shelf space; Toyota Production System
TQM. See Total Quality Management
trainer feedback
training
training hours dilemma
train-the-trainer, TOC
transfer orders (TOs)
transfer pricing
Transition Tree (TRT)
 cluster
 S&T and
 structure of
 using
transportation
transportation lead time
traps, conceptual
Trietsch, D.
TRIZ. See Theory for Inventive Problem Solving
TRR. See time to reliably replenish
TRT. See Transition Tree
trust
Tseng, M. E.
Tsubakitani, S.
Turban, E.
TVC. See totally variable cost
Tyan, J. C.

U

UDE cloud

check/upgrade
consolidating process of
customer
generic cloud and
identify UDE in
parts shortages in
of production manager
solution constructed/communicated in
storyline/building
system

UDEs. See undesirable effects

Umble, E. J.

Umble, M. M.

uncertainty

undesirable effects (UDEs)

cause/effect and
of complex organizations
complex systems with
in core problems
CRT with
dealing with
defining
EC/bank
of health care system
idea/cause-effect relationship with
in solid waste management
systems vs.
symptomatic conflicts and
TP analysis finding

United Methodist Church

unnecessary services

unrefusable offer (URO)

up front buy-in

upstream activity

Uptake Problem

urgency

URO. See unrefusable offer

U-shape

daily problems solved using
for problem-solving
structure of
TOC assumptions and

Uzsoy, R.

V

V-plant research

V-plants

DBR in
product flow in
resource WIP profile of

Vaidyanathan, B. S.

validation

of needs/wants
using CLR

value

customers and
enhancement
lane perspective
metric
perception

Value Focused Management (VFM)

value stream

analysis
map/patients
steps identified in

value-added services (VAS)

Van Slyke, R. M.

variability

- abnormal (red zone)
- of activity duration
- categories of
- closer look at
- convergence points and
- different tools for
- non-critical path and
- in production operations
- resource
- resource contention and
- task
- type of/tools for
- unresolved

variable costs

variance

- expected (green zone)
- fallacy of
- normal (yellow zone)

VAS. See value-added services

VATI analysis

VBP. See Virtual Buffer Penetration

velocity metrics

vendor-managed inventory (VMI)

Ventner, D.

Vermaak, W.

vertical hierarchy

VFM. See Value Focused Management

Viable Systems Model (VSM)

Viable Vision Process (VV)

- case study of
- for health care system
- S&T achieving

vicious cycle

Villforth, R.

virtual buffer concept

Virtual Buffer Penetration (VBP)
visibility
Vision for Successful Dental Practice (Kendall; Wadhwa)
VMI. See vendor-managed inventory
volume exploitation
Von Deylan, L.
VRIO framework
VSM. See Viable Systems Model
VV. See Viable Vision Process
VV S&T tree structure

W

Wadhwa, Gary
Wafer Experiment
waiting
Walker, Ben
Walker, E. D.
Walker, W. T.
Walsh, D. P.
Walton, John
Walton, Sam
Wang, F. K.
Wang, Q. S. G.
Warner, M.
warranty
 CRT
 CS method of
 periods
waste
 current/future gaps of
 Lean and
 of opportunities
 in project environment

Waterfield, N.
Waters, J. A.
Watson, K. J.
WBA. See Why-Because Analysis
WBS. See work breakdown structure
We All Fall Down: Goldratt's Theory of Constraints for Healthcare Systems (Wright, J.; King)
Weakest Link Principle
wedding plan
Weiss, G.
Wesley, John
What to Change
 PSTS and
 Snowflake approach and
 Three-Cloud method
 TOC - Prisons and
white-collar burnout dilemma
Whitman, Walt
Whitney, Eli
Whybark, D. C.
Why-Because Analysis (WBA)
Wiest, J. D.
Williams, D.
Winter, Lamor
win-win
 relationships
 situations
WIP. See work-in-progress
Wolffarth, G.
Womack, James
Wooden, John
Woods, Tiger
work
 behaviors
 ethic, relay runner
 flow, of sales

work breakdown structure (WBS)

work orders

- color coding of

- virtual buffer concept with

workforce

work-in-progress (WIP)

- limiting

- Project S&T reducing

- reducing active

- of resources

workload

Wright, A. C.

Wright, J.

Wu, H. H.

Wu, S.-Y.

X

Xiang, W.

X-Y syndrome

Y

Yeh, M. L.

Yellow Ribbon Project

Yenradee, P.

Yeo, K. T.

“yes, but...”

Young, T.

Z

Zeleny, M.
Zeng, X. L.
zone receipts

1 The OPT scheduling algorithm *Optimized Production Technology* (OPT[®]—a registered trademark of Scheduling Technologies Group Limited, Hounslow, U.K.) was based on the many-body problem in physics.

¹The *APICS Dictionary* (Blackstone 2008, 96) defines Pareto's law as "A concept developed by Vilfredo Pareto, an Italian economist that states that a small percentage of a group accounts for the largest fraction of the impact, value, and so on. In an ABC classification, for example, 20 percent of the inventory items may constitute 80 percent of the inventory value." (© APICS 2008, used by permission, all rights reserved.)

²The *TOCICO Dictionary* (Sullivan et al., 2007, 15) defines *cost-world paradigm* as “The view that a system consists of a series of independent components, and the cost of the system is equal to the summation of the cost of all the sub systems. This view focuses on reducing costs and judges actions/decisions by their local impact. Cost allocation is commonly used to quantify local impact.” In contrast, the *TOCICO Dictionary* (48) defines *throughput-world paradigm* as “The view that a system consists of a series of dependent variables that must work together to achieve the goal and whose ability to do so is limited by some system constraint. The unavoidable conclusion is that system/global improvement is the direct result of improvement at the constraint, and cost allocation is unnecessary and misleading. This paradigm is in conflict with the cost-world paradigm.” (© TOCICO 2007, used by permission, all rights reserved.) For a discussion, see Goldratt (1990).

[3](#)Volume 1 of *The Theory of Constraints Journal* by Eliyahu M. Goldratt and Robert E. Fox (1987) was published.

⁴The distribution solutions are initially mentioned in *It's Not Luck* (Goldratt, 1994) and later in *Necessary but Not Sufficient* (Goldratt, 2000).

⁵Single project critical chain is detailed in *Critical Chain* (Goldratt, 1997).

⁶The Thinking Processes have been utilized in numerous areas besides businesses, including areas such as personal situations (see [Chapter 38](#)), education (see [Chapter 26](#)), and prisons (see [Chapter 27](#)).

[Z](#)The sessions are Operations, Finance and Measurements, Project Management and Engineering, Distribution and Supply Chain, Marketing, Achieving Buy-in and Sales, Managing People, and Strategy and Tactics.

[8](#)The red curve-green curve concept is discussed in detail in Session Eight Strategy and Tactics of the Goldratt Satellite Program.

⁹The S&T trees for each of the five environments (Make-to-Order (Reliable Rapid Response), Make-to-Availability (Consumer Goods), Projects, Retailer, Pay-per-Click) can be accessed and downloaded for viewing with the Harmony viewer at:
http://www.goldrattresearchlabs.com/bin/Harmony_Viewer_0.9.13.5.exe

[10](#) Two of the series (Goldratt, 2008; 2009) have been completed to date.

1It should be noted that when a project is rescheduled because an activity is finished late, the start and finish of the remaining activities are pushed into the future, as is the project completion.

¹The Project Management Institute was founded in 1969 and in the past 40 years has grown into the world's leading project management organization with nearly 500,000 members and credential holders in 180 + countries. <http://www.pmi.org/AboutUs/Pages/default.aspx>, accessed September 5, 2009.

²“The *complexity class* of *decision problems* that are intrinsically harder than those that can be solved by a *nondeterministic Turing machine* in *polynomial time*. When a decision version of a combinatorial *optimization problem* is proven to belong to the class of *NP-complete* problems, which includes well-known problems such as *satisfiability*, *traveling salesman*, the *bin packing problem*, etc., then the optimization version is NP-hard.” (*Algorithms and Theory of Computation Handbook*, 1999, 19-26.) That is, there is no way to identify an optimal solution that includes both a critical path and leveled resources. This fact, however, does not mean that a satisfactory solution cannot be found.

³If you are new to the field of project management, you might find it helpful to review the section in [Chapter 2](#) on the “Development of a Project Network.” That section discusses the use of concepts from the Theory of Constraints to surface potential obstacles to the successful completion of the project. [Chapter 3](#) assumes that all the steps outlined in that section have been accomplished and all activities, including “assumed activities” have been identified.

⁴In traditional project management circles, these terms were developed after the introduction of CC buffers.

⁵These two terms, “path” and “chain,” are used interchangeably.

⁶A matrix structure is one in which people report to more than one superior.

Leveling of resources on a project is now a fairly common practice and such a schedule sometimes is referred to as a resource-constrained critical path schedule in traditional project management circles.

8 Occasionally, a task consumes almost the full amount of task time allocated and the task time therefore should not be reduced (e.g. curing time, bake time, test time).

⁹The *TOCICO Dictionary* (Sullivan, et. al., 2007, 41) defines “*resource buffer—A warning mechanism used in single project environments to ensure that resources working on a Critical Chain task are available when needed.*” (© TOCICO 2007, used by permission all rights reserved.)

[10](#)In terms of the TOC five-step process, acquiring additional resources (Step 7) corresponds to Step 4, “elevate.”

[11](#)This method is a good rule of thumb or heuristic technique, when scheduling manually, but does not always identify the best Critical Chain.

[12](#) Because Resource 4 will proceed to Task F as soon as Task H is completed (following standard CC procedure), there is no need to be overly concerned with timely completion of the top path.

[13](#) Infrequently, a gap may occur on the Critical Chain due to the insertion of a feeding buffer that requires additional resource leveling. These gaps generally are ignored.

[14](#)In this simple example, both Task C and Task I are predecessors to Task J, so the choice of which one on which to focus is debatable. Thus, the situation described here is not typical and Resource 3 may elect to continue working on Task I until it is completed.

[15](#) CC Step 6 corresponds to Step 3, “subordinate,” in the TOC 5-step process.

16In place of resource buffers, some organizations simply report upcoming CC tasks and path starts.

[17](#)CC software will find the best (shortest) schedule, but if scheduling is performed manually, a schedule that “works” is good enough.

[18](#)As noted in footnote 14, both Task I and Task C must be completed prior to the start of Task J and therefore a late completion on Task I might not trigger a move to Task C until Task I is finished.

¹⁹See Guideline XII in [Chapter 2](#).

[20](#)This terminology is used in the *TOCICO Dictionary* and includes a reference to “drum resource” (Sullivan et al., 2007, 7).

[21](#)In the software industry, *integration occurs where various new systems are aggregated and merged together or into older programs.*

[22](#)Note that there may be multiple SRs, reducing required buffer sizes to decouple projects.

[23](#) *Buffer recoveries occur in a similar manner when actual task duration requires less than its estimated (aggressive) duration.*

²⁴With available software, “fever charts” (Newbold, 2008, 112) track buffer consumption and automatically resize buffers throughout the life of a project. An example of such a chart is shown later in [Fig. 3-11](#).

[25](#)For example, homes and businesses located outside of designated flood plains in Atlanta, GA suffered severe flooding in September 2009.

[26](#) There are few instances of budget amounts, once assigned, being returned to the organization.

[27](#)This general statement is not true for *cost-plus* contracts. With cost-plus contracts, the value of the project is actual costs incurred plus some margin, such as 20 percent of total costs. These contracts are becoming quite rare and typically involve research and development type projects where the project deliverable is so unique it is impossible to estimate the total cost to complete the project and achieve its objectives.

[28](#) Because estimation of both overhead costs and driver quantities is imperfect, the allocation rate used during a year (or whatever time until the rate is recomputed) most likely will be adjusted once actual costs and actual driver quantities are known. That means that overhead costs allocated to projects may be adjusted later, sometimes after a project has been completed.

[29](#)If there is uncertainty concerning the actual cost of materials, some amounts representing material cost variability may be added to a project's *budget buffer*, an *account established for task duration variability*.

³⁰A budget buffer represents the budget associated with the safety time removed from individual tasks. The budget associated with the time placed in the project buffer is under the control of the PM while the budget amount associated with the time removed from individual tasks and not connected with the time placed in the project buffer is under the control of the PMO or similar body. However, the budget not under control of the PM can be accessed upon the PM's petition of the PMO.

[31](#) Recall that the *buffer burn rate* is “the rate is calculated as the ratio of the percent of penetration into the project buffer and the percent of completion of the Critical Chain. A buffer burn rate of 1.0 or less is good.” Sullivan et al (2007, 6). (© TOCICO 2007, used by permission all rights reserved.)

[32](#)The Defense Contract Management Agency (2009) of the U. S. Department of Defense has been most active in establishing EVS implementation guidance and project progress metrics.

³³For an extended approach to solidifying change, see Rob Newbold's [Chapter 5](#), "Making Change Stick," in this volume.

[34](#) Similarly, Step 3 is the second layer and so forth until Step 6, which is the fifth layer of resistance.

[1http://www.informs.org](http://www.informs.org). (The Warner Robins Air Logistics Center [WRALC] of the U.S. Air Force won the prestigious Franz Edelman Award 2006 for “Streamlining Aircraft Repair and Overhaul at Warner Robins Air Logistics Center.” Also known as the Super Bowl of Operations Research, it was awarded to WRALC for using Critical Chain to reduce the number of C-5 aircraft undergoing repair and overhaul in the depot from 12 to 7 in just 8 months. The replacement value for these aircraft is estimated at \$2.37 billion.)

[2www.tocico.org](http://www.tocico.org). (The Theory of Constraints International Certification Organization [TOCICO] recognized Boeing Integrated Defense Systems on June 7, 2009. Boeing was presented the North American Achievement Award for its demonstrated longevity in the successful use of Theory of Constraints (TOC) tools and significant contribution to the TOC community. This highly coveted award was handed to Mr. Charles Touns, Boeing Vice President.)

³Enterprise-level means the implementation was not restricted to a single project manager or a small project team, but involved multiple departments.

⁴*Multitasking is shuttling between tasks without finishing either, and hurts the quality of work because people lose concentration.*

⁵*Buffers are unscheduled blocks of time.*

⁶The *TOCICO Dictionary* (Sullivan et al., 2007, 6-7). defines *buffer burn rate* as “The rate at which the project buffer is being consumed in critical chain project management. The rate is calculated as the ratio of the percent of penetration into the project buffer and the percent of completion of the critical chain. A buffer burn rate of 1.0 or less is good. Usage: Some people calculate the burn rate of buffers other than the project buffer. When doing so, use the longest chain of work remaining which feeds into the buffer that is being analyzed. Illustration: If the project buffer is 40% penetrated and the critical chain is only 20% complete, the buffer burn rate is $40/20 = 2.0$. The project manager has a warning that there is a problem, and if it continues, it will possibly jeopardize the project due date.” (© TOCICO 2007, used by permission all rights reserved.)

While managers are an obvious set of stakeholders, depending on the situation, buy-in of an organization's customers and key suppliers might also be needed.

⁸A computer-based learning program, available at www.toc-goldratt.com.

⁹An instructional video by Dr. Eli Goldratt, available at www.toc-goldratt.com.

10 Process of ongoing improvement.

[11](#) Also known as the process for overcoming the five layers of resistance.

[12](#) Maintenance of low WIP is accomplished through Pipeline Planning and Control, which will be discussed later in Step 5.

[13](#) Unlike in high volume production, where the *Drum-Buffer-Rope* (DBR) solution of TOC applies and the drum is a specific resource, in projects the drum is typically a phase.

¹⁴A project plan is different from a *work breakdown structure* (WBS). A project plan is about identifying tasks and precedence relationships among those tasks, whereas a WBS is about subdividing the project into work packages. A task in a project plan can require multiple work packages and vice versa. A project plan is useful for establishing timelines, whereas a WBS might be useful for estimating the total effort.

[15](#)In repetitive environments, these project plans can be stored as templates for future reference.

16It is also known as the *50 percent buffer guideline* because buffers are 50 percent of the sum of task times.

[17](#) Also known as the *student syndrome*; that is, postponing studies until the night before exams.

[18www.realization.com/projectflow/lessons_learned](http://www.realization.com/projectflow/lessons_learned).

[19](#)Sometimes, especially in small organizations, it is not possible to carve out pilots.

20CPI, Cost Performance Index = Budgeted Cost of Work Performed ÷ Actual Cost of Work Performed.

[21](#)SPI, Schedule Performance Index = Budgeted Cost of Work Performed ÷ Budgeted Cost of Work Scheduled.

[1](#)The Cycle of Results and CORE are trademarks of ProChain Solutions, Inc. These marks are used with permission.

²Also called the “*relay racer*” or “*roadrunner*” work ethic (Sullivan et al., 2007, 41).

³Jeannie Duck calls this “the change monster,” saying that one needs great determination to see change efforts through to the point where the needed changes have become the standard way of doing things (Duck, 2001, Part 5).

⁴This includes several Fortune 200 companies across dozens of business units. Some results for Fortune 200 companies are described in Newbold (2008, 5). The logic is described in the following pages.

⁵There are many references on Current Reality Trees; see, for example, Scheinkopf (2000, [Chapter 8](#)).

[6](#)We have seen cases where pockets of Critical Chain use persist in large organizations. Typically, a few enthusiasts are able to derive enough benefits and win enough converts to offset the overall lack of momentum. This is a frustrating road to travel because the benefits are so clearly less than what is possible.

[↗](#) See again Hobbs and Aubry (2006), discussed previously.

⁸Cialdini (1993, [Chapter 3](#)) notes that expressions of support make people more likely to actually give it later on.

⁹For more discussion of measurements, see Newbold (2008, [Chapter 12](#)).

[10](#)This is the standard Six Sigma improvement process, for which there are many references.

[11](#) See Covey (1989, 306).

[12](#)This TOC concept is mentioned in Sullivan et al. (2007, 30). Also, look under *buy-in*. A comparison of CORE with the Layers of Resistance may be of special value for TOC practitioners because true “buy-in” requires trust, which requires feedback. However, be warned that there are many incarnations of the Layers of Resistance.

[13](#)This is the combat operations process developed by John Boyd of the U.S. Air Force. For an excellent description, see Richards (2004).

[14](#) See Scheinkopf (2000) and Dettmer (2007).

[15](#) See Goldratt et al., (2002).

[16](#)For a more complete discussion, see Newbold (2008, 172).

¹⁷[Table 5-4](#) contains the Solution Selling steps for dealing with “latent opportunities,” in which the customer is not actively looking for a solution. For the full process, see Eades (2004, 38-41).

[18](#)PDCA is also known as the Deming cycle or the Shewhart cycle. See Deming (1982, 88-89).

¹⁹For an in-depth discussion of Critical Chain pilots, see Newbold (2008, [Chapter 17](#)).

²⁰For one of the earliest discussions on the five focusing steps, see Goldratt (1990, [Chapter 1](#)). The associated loops are shown clearly in Newbold (1998, 150).

²¹See, for example, <http://billiondollarsolution.com/blog/?p=70>, accessed July 12, 2009.

[22](#) See Lewin's 1947 paper "Frontiers in Group Dynamics" as reprinted in Lewin (1997, 330).

[23](#)For much more on this subject see Kotter (1996; 2008).

²Creating significant change in a traditional body of knowledge is known as a paradigm shift and generally encounters significant resistance. TOC is just such a change and attacks the very foundation of traditional business knowledge and practice. Goldratt (2003b) describes his struggle to improve production.

³A registered trademark of Scheduling Technologies Group Limited, Hounslow, UK.

[4](#)© E.M. Goldratt used by permission, all rights reserved.

5 Some researchers combine protective and excess capacity into what they call *idle capacity*. This is capacity that when operations are running smoothly, the resource is idle. For the manager, the problem with idle capacity is how much is protective and how much can be trimmed?

⁶In TOC, *starvation* is measured only at the constraint and *exists when the constraint is idle caused by lack of material*. In contrast, *blocking* occurs when *the constraint doesn't have space to offload finished units and therefore must sit idle until space is freed up*.

Over the years, Goldratt and others used computer animated simulators to teach the concepts of DBR scheduling and Buffer Management in V, A, and T environments. The early versions of the teaching simulators were DOS based; a newer Windows-based version is provided in Goldratt (2003).

8 Material protecting the constraint in this situation is technically raw material and not considered WIP until released to the line.

⁹Most implementations today recognize two other regions: a *Black Region* that *indentifies orders that should have been completed and are now late*, and a *White Region* that *identifies orders that should not have been released but were released early*.

[10](#)Goldratt (2009) provides an insightful comparison of Henry Ford's assembly line, Dr. Ohno's Toyota Production System and his Drum-Buffer-Rope system.

¹The *APICS Dictionary* (Blackstone, 2007, 86) defines *nervousness* as “*The characteristic in an MRP system when minor changes in higher level (e.g., level 0 or 1) records or the master production schedule cause significant timing or quantity changes in lower level (e.g., level 5 or 6) schedules and orders.*” (© APICS 2008, used by permission, all rights reserved.)

Copyright © by Mokshagundam (Shri) Srikanth.

²Senge (1990, 71) defines *detail complexity* as “the sort of complexity where there are many variables” and *dynamic complexity* as “situations where cause and effect are subtle, and where the effects over time of interventions are not obvious.”

³TPS uses a *two-card kanban system*. The *APICS Dictionary* (Blackstone, 2008, 142) defines this as “(a) *kanban system where a move card and production card are employed. The move card authorizes the movement of a specific number of parts from a source to a point of use. The move card is attached to the standard container of parts during movement to the point of use of the parts. The production card authorizes the production of a given number of parts for use or replenishment.*” (© APICS 2008, used by permission, all rights reserved.)

[4](#)© TOCICO 2007, used by permission all rights reserved.

⁵For a discussion of these measures, see [Chapter 13](#) of this Handbook.

⁶In TOC, consumer products are managed with MTA, a pull supply chain system, where traditional supply chains use a make-to-stock (MTS) system (min-max or reorder point/economic order quantity). See [Chapters 10, 11, and 12](#) of this Handbook.

[↑](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 48) defines *time buffer* as “*Protection against uncertainty that takes the form of time.*” © TOCICO 2007, used by permission, all rights reserved.

⁸In TOC terminology in the *TOCICO Dictionary* (Sullivan et al. 2007, 13) each is called a *control point*, which is defined as “(a) *key point in the flow of work through an operations environment that, if not managed properly, has a high probability of decreasing due date performance. Control points include gating operations, convergent points, divergent points, constraints, and shipping points. Usage: In TOC operations management, sequencing schedules at the control points to match the drum schedule and/or shipping schedule increases the probability of on-time performance.*” (© TOCICO 2007, used by permission, all rights reserved.)

[9](#)TOCICO 2007, used by permission, all rights reserved.

[10](#) © E. M. Goldratt used by permission, all rights reserved. The example is taken from Goldratt (2003). This work comes with a CD, which has this as well as other examples for readers to develop their own production schedules and see the results through simulation.

11 This is called a “*free product*” as no additional direct labor expense is required to produce it.

¹²The *APICS Dictionary* (Blackstone, 2007, 108) uses a similar term: *product structure*—“the sequence of operations that components follow during their manufacture into a product. A typical product structure would show raw material converted into fabricated components, components put together to make subassemblies, subassemblies going into assemblies, and so forth.” (© APICS 2008, used by permission, all rights reserved.)

¹³The *TOCICO Dictionary* (Sullivan et al., 2007, 15) defines *cost-world paradigm* as “The view that a system consists of a series of independent components and the cost of the system is equal to the summation of the cost of all the sub systems. This view focuses on reducing costs and judges actions/decisions by their local impact. Cost allocation is commonly used to quantify local impact.” (© TOCICO 2007, used by permission all rights reserved.)

¹See [Chapter 10](#) of this Handbook.

Copyright © 2010 by Eli Schragenheim.

[2](#)For the interested reader, see the nine OPT[®] rules in Goldratt and Fox (1986, 179).

[3](#) © E. M. Goldratt used by permission, all rights reserved.

⁴For a more comprehensive discussion of the OPT[®] software, see Fry, Cox, and Blackstone (1992).

⁵The more elaborate DBR methodology, like the ones described in the last part of *The Haystack Syndrome* (Goldratt, 1990a) and in the Disaster software developed at the end of the 1980s, included an algorithm to identify and plan several capacity constraints in detail. Still, having to live with interactive capacity constraints was viewed as causing the system to be unstable and the direct recommendation was to elevate the interactive constraints, so only one capacity constraint would remain.

6 Technically the operations on divergence operations, where the parts could go to different end products, were given a schedule to prevent stealing. However, the capacity of these resources was not checked and the other operations of those resources were not scheduled.

[I](#)© E. M. Goldratt used by permission, all rights reserved.

⁸Green-yellow-red buffer regions are the backbone of BM. This topic is mentioned in [Chapter 8](#), and the assumption we adopt here is that the reader is familiar with the basic concept.

⁹In scheduling the CCR and especially when trying to save setups on the CCR, it could be the case that an order whose due date is later in time would be scheduled earlier to save CCR setup time.

10Sometimes different families of products are going through different work centers and thus it could be the case that one family has a specific resource whose load control dictates the planning, while another family has a different resource where load control is required.

11In all situations, the scheduler should ensure that *all materials, specs, tooling, etc., are available prior to release of the manufacturing order to the shop floor (full kitting)*.

[12](#)*In DBR, the production buffer would be equal to the constraint buffer plus the shipping buffer. In simplified DBR, the production buffer would be equal to the shipping buffer.*

[13](#) Other actions such as overlapping processes can also be implemented to reduce lead time significantly.

14 There are several exceptions to the rule. Most notably are shops where a dedicated assembly line dominates production. These types of “I”-plants have naturally very low WIP within the line itself.

[15](#)BM is also the source for focusing the efforts on balancing the flow. This is mentioned later in this chapter as POOGI—a *process for ongoing improvement*.

¹⁶Different families of products could have considerably different production buffers. We do assume that using a different buffer size is done *only* when the buffer size is at least 25 percent more or less than buffer sizes already defined.

17The reason is that most orders are found close to the end of the production process when the customer complains. Top management then increases the capacity of these operations, thinking these operations are the cause of the problem of lateness. Thus, normally the last operations have ample protective capacity.

18 Actually, this perception is wrong because the processing time is a negligible part of the lead time and the main part dictating the delivery time is the required waiting for the resource with the least capacity.

¹⁹The *APICS Dictionary* (Blackstone, 2008, 56) defines *funnel experiment*— “An experiment that demonstrates the effects of tampering. Marbles are dropped through a funnel in an attempt to hit a flat-surfaced target below. The experiment shows that adjusting a stable process to compensate for an undesirable result or an extraordinarily good result will produce output that is worse than if the process had been left alone. “(© APICS 2008, used by permission, all rights reserved.)

[20](#)A detailed discussion on how to deal with sequence-dependent-setups can be found in Schragenheim, Dettmer, and Patterson (2009, 79-86).

[21](#) *When orders are late we sometimes call them "black."* The author believes that black is not part of the priority system and not necessarily a black order has higher priority over red orders.

[22](#)For a more detailed discussion, see Schragenheim et al. (2009, 74-79).

[23](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 27) defines “*I-plant—A production environment where materials generally flow through a direct sequence of operations. The logical flow of materials resembles the letter I in the sense that is, there are few divergent points, as in a V-plant, and few convergent points, as in an A-plant. Examples: Transfer or assembly lines such as used to assemble lawn mowers.*” (© TOCICO 2007, used by permission, all rights reserved.)

¹For those interested, marketing is covered in [Chapter 22](#) and sales management is covered in [Chapter 23](#) of this Handbook.

²For those unfamiliar with MRP see, for example, Arnold, Chapman, and Clive (2008).

³In the MRP literature, this condition is called *nervousness*, which the *APICS Dictionary* (Blackstone, 2008, 86) defines as “(t)he characteristic in an MRP system when minor changes in higher level (e.g., level 0 or 1) records or the master production schedule cause significant timing or quantity changes in lower level (e.g., level 5 or 6) schedules and orders.” (© APICS 2008, used by permission, all rights reserved.)

⁴The difference we mean here between sales and consumption is when certain items “vanish,” either because they were scrapped, stolen, or lost. The default action here is still to replenish them.

⁵The *shipping buffer* is a *time buffer* used in *DBR* to protect the *due date* of an order. In *MTA*, we need to protect the availability, so it has to be a different type of a buffer but for the same purpose of protecting the satisfaction of the clients. Thus, the quotation marks mean that it is not the same but the objective is similar.

6 Replenishment time in TOC differs considerably from traditional inventory management. In both the min-max and reorder point/economic order quantity inventory system, items are sold over some time period and only when the minimum point or the reorder point is reached and an order is placed. Replenishment time starts when an order is placed for an item and ends when the item is restocked and available for sale. Note in the TOC, *replenishment is triggered by time (daily or weekly maybe)* and in traditional methods triggered by an inventory level falling to or below some reorder level.

[Z](#)Note that the term “reliable replenishment time” is different than the term “replenishment time,” which is the average.

⁸The author defines the abstract term of “*protective capacity*” as *the level where the lack of immediate available capacity starts to cause real damage*.

⁹The regular planned load is the one for daily use. A full planned load will be defined later as consisting of all required replenishments including those that were not released.

10 It could be the case where the minimum batch is aimed to keep another resource, with long setups but fast processing pace, from becoming a bottleneck. Therefore, the minimum batch is a must, but once this is under control the capacity control is focused on the CCR.

[11](#)Deming devised an experiment to show the results of such tampering before a process has stabilized. The *funnel experiment* in the *APICS Dictionary* (Blackstone, 2008, 56) is defined as “An experiment that demonstrates the effects of tampering. Marbles are dropped through a funnel in an attempt to hit a flat-surfaced target below. The experiment shows that adjusting a stable process to compensate for an undesirable result or an extraordinarily good result will produce output that is worse than if the process had been left alone.” Here you are adjusting again before the process has stabilized. (© APICS 2008, used by permission, all rights reserved.)

[12](#)Of course, we assume here that all the efforts to exploit the current capacity of the CCR have been done. Actions like staggering breaks and lunches plus overlapping shifts reduce or eliminate down time on the constraint and increase capacity by over 10 percent.

¹³Goldratt (2009) compares and contrasts the traditional assembly line, the Toyota Production System, and drum-buffer-rope. Based on his analysis, Goldratt provides four concepts of supply chains. These four concepts were discussed in [Chapter 9](#). The fourth concept, flow, relates heavily to MTA.

¹⁴The *APICS Dictionary* (Blackstone, 2008, 144) defines *vendor-managed inventory (VMI)* as “(a) means of optimizing supply chain performance in which the supplier has access to the customer’s inventory data and is responsible for maintaining the inventory level required by the customer. . .” (© APICS 2008, used by permission, all rights reserved.)

[15](#) Forecasting the demand means manually changing the target levels according to the “maximum” forecast within the reliable replenishment time and doing it before the peak is supposed to start.

[16](#)This problem is not unique to TOC. Manufacturers have long been trying to identify hybrid production systems that would work under the push system. The *APICS Dictionary* (Blackstone, 2008, 61) defines *hybrid production method* as “(a) production planning method that combines the aspects of both the chase and level production planning methods.” (© APICS 2008, used by permission, all rights reserved.)

17When we speak about a dedicated production line, then the production lead time could be very fast and the actual replenishment time is the wait time for the line itself.

¹⁸See [Chapters 16](#) and [20](#) in this volume.

¹Editors' note: This chapter describes a state-of-the art software package and how it addresses the realities of complex manufacturing and supply chain environments. The author was invited to contribute to the handbook as he has studied these environments using the TOC Thinking Processes, has an in-depth understanding of the causal linkages from symptoms to core problem(s) and has struggled with developing comprehensive solutions to these core problems. He is an expert in this development work.

²In distribution, the *APICS Dictionary* (Blackstone 2008, 131) makes a significant distinction between the meanings of these two terms. A *product is a good within the supply chain*, while an *SKU* is defined as: “*an item at a particular geographic location.*” © APICS 2008, used by permission, all rights reserved.

³Sales at the end of the chain in the long-term are the only measures that matter. If a sale is made within the links of the supply chain, it is to fill inventory positions for future sales at the end of the chain. Therefore, once the supply chain is filled it will sit without movement until the consumer purchases at the consumption point. Most managers within a traditional supply chain focus only on sales to the next link, not sales at the end of the chain.

Copyright © 2010 by Amir Schragenheim.

⁴In TOC terminology, we call these *undesirable effects* (UDEs) and we search for the core problem causing these UDEs.

⁵Throughout this chapter I will use the term manufacturer/distributor to represent either a supply chain based on a company that manufactures the majority of parts that flow through its supply chain or a company that purchases parts from one or several manufacturers and distributes them through its supply chain.

[6](#)The Z-value from a normal distribution table for .05 in the right tail is 1.645.

[7](#)A wide array of forecasting methods exists for modeling trend, seasonal, cyclic, and random factors, and their combinations at the aggregate level but all perform poorly at the consumption point.

8 The sudden death of Michael Jackson caused the sales of his CDs to stockout most of the complete supply chain of all of his recordings. The same phenomenon occurred with Elvis Presley's death.

⁹In most push systems, since the finished goods inventory is predominantly at the retailer end of the chain, manufacturing warehouses are generally small and hold only a few days of inventories. In a pull system, a larger warehouse is generally needed as inventory is held higher in the chain at the source.

10The supply factor is usually ignored in tactical and strategic decision making. Most efforts for improvement are directed at the demand side, especially trying to come up with more sophisticated forecast algorithms. Most new ERP systems developers keep coming up with new forecasting algorithms, neglecting the supply side completely.

11The replenishment part of the TOC pull distribution solution can be compared with a min-max model of replenishment where the min equals the max.

[12](#) group the two methods together even though there are some minor differences. If we set the difference between the min and max values of the min-max model at the EOQ value then the results are quite similar. In traditional management this min-max gap is often set at the EOQ. Under these conditions we get very similar graphs. The two systems will give different results only when the EOQ is quite small or, alternatively, the consumption is chaotic rather than continuous.

[13](#) Despite the fact that the TOC solution discourages using forecasts, sometimes (like in the case of seasonality discussed later in the chapter) forecasts are still needed.

14The use of S-DBR, The TOC methodology for managing production in supply chains is covered in detail in [Chapter 9](#) of this volume.

¹⁵The *TOCICO Dictionary* (Sullivan et al., 2008, 49) defines *totally variable cost (TVC)* as “Those costs that vary 1-to-1 for every increase in the number of units produced.” (© TOCICO 2007, used by permission, all rights reserved.)

16In some extreme cases, it is better to use different sizes of zones especially when the replenishment time is more than three months or, alternatively, the products are short shelf-life products.

[17](#)Note that these relate to all types of consumption points—shops, regional warehouses (RWHs), plant warehouse (PWH), central warehouse (CWH), raw materials warehouse (RMWH), etc.

[18](http://www.inherentsimplicity.com/) See the Inherent Simplicity Web site for an example and complete discussion of this concept at: URL=<http://www.inherentsimplicity.com/>.

19 Too long is a term that depends on the specific environment—typically one to two weeks.

[20](#) It is important to note that this categorization is done after the TOC solution has been put in place. Therefore, the stock levels are already adjusted based on reality and not on a managerial decision.

[21](#)We use the number of turns here in a relative sense. If one did a Pareto analysis of retailer items based on inventory turns and ranked the items by number of turns, the A items (cheetahs) would have a high number of turns relatively speaking, the B items (regularly running) would have a moderate (around the mean) number of turns, and the C items (elephants) would have a low number of turns.

[22](#)This conclusion also stems from analyzing the sales results obtained from TOC implementations.

²³The *APICS Dictionary* (Blackstone, 2008, 67) defines *inventory turnover* as “(t) he number of times that an inventory cycles, or ‘turns over,’ during the year. A frequently used method to compute inventory turnover is to divide the average inventory level into the annual cost of sales. For example, an average inventory of \$3 million divided into an annual cost of sales of \$21 million means that inventory turned over seven times” The traditional definition for *replenishment lead time* in the *APICS Dictionary* (Blackstone, 2008, 117) is “(t) he total period of time that elapses from the moment it is determined that a product should be reordered until the product is back on the shelf available for use.” (© APICS 2008, used by permission, all rights reserved.)

²⁴Of course, there are always exceptions to any rule, so judiciously choose candidates you want to eliminate. Maybe you get a low margin on bread (necessity goods), but a grocery store that doesn't stock it might lose a lot of business for all other items. Complementary goods are another class of goods that require scrutiny in that one may have a high margin and another a low margin (toothpaste and toothbrush, for example).

²⁵A year is generally used but for seasonal or fashion goods, the length of time must be modified to fit the situation.

[26](#)If severe fluctuations exist, a larger fluctuation factor should be used but beware of demand patterns in an environment where volume discounts are given. These discounts distort the demand pattern significantly.

[27](http://www.inherentsimplicity.com/) See the Inherent Simplicity's Web site for references at:
<http://www.inherentsimplicity.com/>.

[28](#)For example, think about the change in demand for beer consumption for the summer versus the winter; for a national holiday weekend across the country versus a regular weekend; for a home football game weekend for a major college town versus a regular weekend in that same town.

[29](#)Inherent Simplicity is a leading provider of TOC distribution/replenishment software, and has faced and solved these situations based on client needs. I will discuss several different scenarios that exist in reality and the solution methods derived for a pull system by this company. This is not meant as an endorsement of the software but to illustrate a practical way of addressing that issue if it exists in your environment.

[30](#)A forecast can be as simple as noticing that retail sales on the weekend of a college football game increase six times over a normal session weekend. By determining when the next season's home football games are, one can plan to have inventories to cover these peaks for the season.

[31](#)The Inherent Simplicity methodology involves a series of steps to be performed for each SDC and for each SKU group separately.

[32](#)BM priority during this time might be skewed as well. Suppose the buffer was 100 and now it increases to 400. Even if we had a full buffer at the site, the VBP at the site would now show a 75 percent penetration—quite deep in the red.

[33](#)It is imperative that management judgment is used to determine if there is a decrease in demand immediately following the SDC and therefore make adjustments for the SDC impact. An advertising campaign based on price reduction for a cereal is an example of a sales decrease. There are also situations where the normal demand would resume almost immediately. For example, while retail sales are high (SDC) for a football game weekend (caused by out-of-town consumers), normal demand resumes almost immediately.

[34](#)If you know the number of SKUs in your system, then you can skip this step; it just shows the calculations of converting items by locations to SKUs.

[35](#)A supply chain board game comprised of three products, six retailers, two RWHs, and one PWH has been used to teach the differences between managing under the traditional push systems (ROP/EOQ and min-max) and the TOC distribution/replenishment system. It is described in Cox and Walker (2006).

³⁶See “The Science of Successful TOC Holistic Implementation” presented by Mickey Granot at the TOCICO 2008. For further references of Inherent Simplicity, please refer to Inherent Simplicity’s Web site at: <http://www.inherentsimplicity.com/>.

[37](#)As in all business functions, distribution goes through cycles of centralization for cost control to decentralization for flexibility/responsiveness. CWH has been tried using distribution requirements planning and distribution resource planning (traditional push systems) and failed. The move then was to decentralization. Now organizations are switching back to centralization with ERP and supply chain software.

¹When considering an APS, an understanding of the materials of this chapter and other chapters in this section is necessary.

¹An excellent summary may be found in Johnson and Kaplan (1987, [Chapter 2](#)).

²For a critical review of textbook treatment of Frederick Taylor's principles and the Hawthorne experiments by Elton Mayo and others, see Whitehead (1938) and Olson et al. (2004).

³A long-time acquaintance, a senior partner in an international audit firm who had majored in accounting during his undergraduate years, confessed to the author that he would recommend that his children major in finance, not accounting.

⁴See Peters and Waterman (1982, 34-39) for a discussion of this situation and relevant references.

⁵See especially Chapters 6 and 8 of Johnson and Kaplan.

⁶Also sometimes called “margin” costing or “contribution margin” costing since the first subtotal on this type of statement is called “contribution margin” or “gross margin.”

[Z](#)Dopuch and Birnberg (1969, 472) relate the original internal flow concern to a 1936 article (Harris, 1936).

8 Time-driven allocations later were promoted in an attempt to overcome some of these deficiencies (Kaplan and Anderson, 2003; Everaert and Bruggeman, 2007), but were generally unsuccessful (Cardinaels and Labro, 2009).

⁹While Kaplan and Norton generally are credited with the development of the balanced scorecard, their renowned book states in the preface that a dozen companies met bi-monthly throughout 1990 to develop a new model (Kaplan and Norton, 1996, vii).

[10](#) Many dissertations have been based on balanced scorecard concepts. For example, Deem (2009) conducted a study that found a positive relationship of balanced scorecard effectiveness and organizational culture.

[11](#)This is not for lack of improvement opportunities in accounting. Simply reducing the time typically required to close the books would be of great benefit.

[12](#) Kaizen blitz^(SM) is a service mark term of the Association of Manufacturing Excellence.

[13](#)An extensive example later in the chapter shows the unintended negative effect of inventory reduction on the income statement.

[14](#)The vast majority of recent budget research concerns governmental budgeting, which has unique issues (Kelly and Rivenbark, 2008) and will not be addressed here.

¹⁵Due to space constraints, most textbooks illustrate quarterly budgets (Hilton, 2009, [Chapter 9](#); Garrison et al., 2010, [Chapter 9](#)), but companies need budgets prepared on at least a monthly basis in order to accurately predict cash needs and when it might be possible to invest extra cash.

16It is the author's experience that marketing people make the most sophisticated presentations.

[17](#)If an organization has more than one constraint, the typical accounting recommendation is to use linear programming to find the best product mix. This material is frequently skipped by accounting professors.

[18](#)The 80 units of Product Z would require 400 minutes on Resource 2, the 90 units of Product X would use 1800 minutes, and 10 units of Product Y would use the final 200 minutes. In reality, capacity usage would not be scheduled at 100 percent, but a lower capacity availability for all resources would not change the essential results presented here.

19 For example, labor allocated according to minutes spent on each product; fixed manufacturing overhead allocated based on total variable manufacturing costs, only whole units may be sold, etc.

[20](#) Located on the Web at:
www.mhprofessional.com/TOCHandbook.

[21](#) Three pools, Planning, Processing, and Support, along with individual drivers, are used to allocate all costs to products and arrive at total product profit per unit as shown in cells G128-J135 of the “Throughput_Examples” spreadsheet located at: www.mhprofessional.com/TOCHandbook.

[22](#) Revenues minus variable cost of sales, minus other variable expenses (general, selling, and administrative), to show Throughput (contribution margin or gross margin), minus Operating Expenses (fixed costs) to arrive at operating income before income tax expense.

[23](#) See a complete illustration of the Throughput format in the second spreadsheet of the file entitled “InventoryReductionExample” located at www.mhprofessional.com/TOCHandbook.

[24](#) See the section on direct or variable costing income statements, which contain all the elements of a Throughput income statement.

[25](#) Only 10 units of Product Y are planned for production, therefore $10 \times .07 = 0.7$.

[26](#)Total weekly wages of \$4800, divided by 9600 (40 hr × 60 min × 4 resources) total labor minutes. This is the actual cost of labor per minute; the traditional applied rate of \$0.53333 is based on expected production matching demand which requires 9000 labor minutes.

[27](#)This problem originally was pointed out in Goldratt (1990, 97-99).

[28](#)[\$183 throughput (contribution margin) for Product Y] × 2
units = \$366.

[29](#)Note that many companies using traditional costing also (incorrectly) would include savings in fixed overhead costs that will merely be transferred to other products.

[30](#) See the “Total Time Used/Week” schedule—P27: U33—in the *Throughput Example* file located at www.mhprofessional.com/TOCHandbook.

³¹See Throughput_Examples spreadsheet located at www.mhprofessional.com/TOCHandbook, cells CH1: CO40.

[32](#)Accountants customarily carry costs to four significant digits to the right of the decimal so that aggregated totals from computing costs for many units will be more precise. For a more in-depth look at this issue, see Eden and Ronen (2007).

[33](#) $\$40 - \$29.97 = \$10$; $\$10 \times 100$ units (90 for Product X and +10 for Product Y) = \$1000.

[34](#)90 units for Product X and 50 units for Product Y.

³⁵See the Throughput_Examples spreadsheet located at www.mhprofessional.com/TOCHandbook, cells CQ1: CZ40.

³⁶Operating income would drop from \$12,858 to \$7,858, a decrease of \$5,000, the cost of the change. See the *Throughput_Examples* spreadsheet, cells BV47:CF82 at www.mhprofessional.com/TOCHandbook.

[37](#) See the Throughput_Examples spreadsheet, located at www.mhprofessional.com/TOCHandbook, cells CH47:CP82.

[38](#) While recent cost and management accounting texts acknowledge TOC, attempt to define it, and recognize its connection to the contribution-margin-per-unit-of-constraint decision, they do not address the impact of a constraint on numerous other operating decisions such as make versus buy, adding or dropping product lines, or special orders.

[39](#) Boeing has used this approach upon occasion to clean up its WIP inventory (Henkoff, 1998; Skapinker, 1998).

[40](#) Since it is merely a timing issue, if inventory ever increases back to original amounts the income discrepancy will disappear.

[41](#)Quantities of materials held should be a function of (1) frequency of supplier delivery, (2) company ability to reliably predict consumption levels during supplier lead time, and (3) vendor reliability in meeting promised shipment dates and in meeting specified quality (Goldratt, 1990, 108).

[42](#) Many TOC practitioners simply use Throughput/Inventory, but the above formula more correctly matches the traditional definition of inventory turns.

[43](#) Unavailability of credit was experienced by many businesses during the recession of late 2008, 2009, and the first part of 2010.

[44](#)Find at:
<http://www.nasba.org/nasbaweb/NASBAWeb.nsf/FNAL/CandidateBulletin?opendocument>. Business Environment and Content requirements begin on p. 29. Accessed August 29, 2009. (Section VI, C, 5, p. 33, states that candidate responsibilities include “Management philosophies and techniques for performance improvement such as *Just-in-time* (JIT), Quality, Lean, Demand Flow, Theory of Constraints, and Six Sigma.”)

[45](#)The Certified Management Accountant examination sponsored by the Institute of Management Accountants has also updated its Content Specification Outline to include TOC and TA, effective May 1, 2010.

[46](#) Numerous threads to this effect can be found in the archives of the AECM Web site (Jenson et al., 2009).

⁴⁷The seven FASB Concepts Statements may be found at the following Web site:
<http://www.fasb.org/jsp/FASB/Page/SectionPage&cid=1176156317989>, (accessed March 20, 2010.)

⁴⁸See [Chapter 14](#) in reference to dysfunctional behavior caused by local measures.

¹Note: Recent events in the world economy have made companies hypersensitive to cash. Cash and cash flow are necessary conditions of doing business and should not be treated as objectives in and of themselves. In a good TOC system, however, decisions tend to be filtered by their total cash implications since the metric system assigns product cost based on the direct cost method instead of a standard product cost basis that would include both fixed and variable overhead.

[Chapter 15](#) provides a simple numerical example illustrating these points.

[3](#) Johnson and Kaplan (1987) provide a history of managerial accounting and describe many of the problems created by its use.

⁴The *APICS Dictionary* (Blackstone, 2008, 14) defines a *break-even chart* as “(a) graphical tool showing the total variable cost and fixed cost curve along with the total revenue curve. The point of intersection is defined as the *break-even point* (i.e., the point at which total revenues exactly equal total costs)” (© APICS 2008, used by permission, all rights reserved.) This definition uses the traditional view of fixed and variable costs. This difference between TOC and traditional accounting creates vast differences in decision making in most situations. TOC falls in line with fundamental economics, therefore giving the correct answer.

⁵The *TOCICO Dictionary* (Sullivan et al., 2007, 40) defines *protective capacity*—“Resource capacity needed to protect the throughput of the system by ensuring that some capacity above the capacity required to exploit the constraint is available to catch up when disruptions inevitably occur. Non-constraint resources need protective capacity to rebuild the bank in front of the constraint or capacity constrained resource (CCR) and/or on the shipping dock before throughput is lost.” (© TOCICO 2007, used by permission, all rights reserved.)

⁶For an in-depth explanation and case study on pricing indifference modeling, we refer you to [Chapter 9](#) of *The Measurement Nightmare, How the Theory of Constraints Can Resolve Conflicting Strategies, Policies, and Measures*, by Debra Smith, St. Lucie Press, 2000.

[↑](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 40) defines *time buffer*—“*Protection against uncertainty that takes the form of time.*” (© TOCICO 2007, used by permission, all rights reserved.)

⁸The *TOCICO Dictionary* (Sullivan et al., 2007, 43) defines *stock buffer*—“*A quantity of physical inventory held in the system to protect the system’s throughput.*” (© TOCICO 2007, used by permission, all rights reserved.)

⁹*Capacity buffer is the sprint or protective capacity placed at non-constraint resources to protect against Murphy.*

[10](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 41-42) defines *roadrunner work ethic* as “(t)he work rules in the drum-buffer-rope or critical chain project management (CCPM) systems. The rules are: if there is work available start it immediately; if there is more than one work-order/task in queue choose the one with the highest system-priority; work at full speed without stopping until the work is completed; produce zero defects and pass the work on immediately; if there is no work available stay idle.” (© TOCICO 2007, used by permission, all rights reserved.)

¹¹See [Chapter 13](#).

[12](#) Whether the environment is CCPM, DBR, Replenishment, or ASR, the specifics and the supporting tools of reporting and measurement, direct the necessary change but the five questions remain the same.

¹³The *TOCICO Dictionary* (Sullivan et al., 2007, 48) defines *time buffer*—“Protection against uncertainty that takes the form of time. See: *assembly buffer, buffer, drum-buffer-rope, drum buffer, capacity buffer, feeding buffer, project buffer, shipping buffer.*” (© TOCICO 2007, used by permission, all rights reserved.)

¹⁴See [Chapter 24](#) for an explanation of these processes.

¹© E. M. Goldratt used by permission, all rights reserved.

²Thomas Aquinas (1225–1274), the most important Catholic medieval philosopher and theologian, was most likely the one who came up with the classification of “sins of omission” and “sins of commission,” but there are references to “the sin of failing to do something good when you know you should” and “the sin of doing the wrong thing” in the Bible (the Good Samaritan parable and Ten Commandments are classic examples of such references).

³The improvement challenges identified by Barnard are similar in nature to those identified by Dr. Eli Goldratt in *The Choice* (2009, 157-158) as obstacles that must be overcome to achieve a full life (through the choice to think like a scientist). These include “perception that reality is complex,” “accepting conflict as given,” “blaming others,” and “thinking that you know.”

⁴See [Chapter 25](#) of this Handbook.

⁵See [Chapter 16](#) of this Handbook for a case study involving such a five-day workshop.

[6](#) © E. M. Goldratt used by permission, all rights reserved.

1 A time buffer is a release control mechanism that protects the due date of an order against expected disruptions (Murphy) by releasing it significantly earlier than the actual processing time of the order while not so early that it will contribute to long queues, high WIP and longer lead times. The general rule in TOC is that time buffers are set at 50% of the pre-TOC implementation lead time and are divided into three equal zones (green, yellow and red) with orders that enter red status, triggering expediting actions.

⁸Goldratt's latest insights on the required focusing mechanism for continuously improving Operations, Distribution, Projects, and Sales can be found in Goldratt's generic S&T that have been released into the public domain. These can be found in the S&T Library embedded in HARMONY (S&T Expert System) downloadable from www.goldrattresearchlabs.com.

⁹TOC draws a distinction between issues that are truly important and issues that are of a peripheral nature. Goldratt has used a wonderful word from the Israeli slang to describe trivial issues: choopchick. Apparently, “choopchick” comes from the behavior of teenage boys that spend so much time on their “choop”—their hair—with a real belief that this will help them attract “chicks”—teenage girls—therefore, it refers to an action that takes significant efforts, but accomplishes very little else. The effect of choopchicks within the management process can be devastating. Wasting time and resources on relatively unimportant issues diverts efforts from genuinely significant concerns.

[10](#) See, for example, T. Pyzdek and P. A. Keller, 2010. *The Six Sigma Handbook*. 3rd ed., New York, McGraw-Hill, 127-133.

[11](#) DMAIC is the acronym used for the Six Sigma process that includes five steps—Define, Measure, Analyze, Improve, and Control.

[12](#)All the latest S&Ts are available at www.goldrattresearchlabs.com with a free download of Harmony S&T viewer.

¹Examples of TOC success stories that were not sustained, or eventually failed due to one or more of the listed mistakes include Bethlehem Steel and General Motors.

²Dr. Eli Goldratt (Goldratt, 1999b) calls the exponential growth in performance the “red curve” and the improved stability curve with diminishing returns towards a maximum level, the “green curve.”

³(Cost World: Managing using rules that assume $\text{Global Optima} = \text{Sum of Local Optima}$). Throughput World: Managing using rules that acknowledge $\text{Global Optima} \neq \text{Sum of Local Optima}$. Global optima is achieved through better protecting, exploiting or elevating system constraints to increase System goal units (Throughput) with the minimum Operating Expenses and Investment.

⁴Goldratt (2009) has recently shared his own realization that TOC can be summarized in one word—FOCUS, as outlined in [Chapter 1](#). Goldratt also explains that for him the definition of FOCUS is simply preventing the two mistakes identified by Ackoff (2006). Real FOCUS means *doing what should be done and not doing what should not be done*.

⁵A large bank of references of TOC success stories can be found at www.toc-goldratt.com.

⁶The exception to the rule are the holistic implementations of TOC within the U.S. Navy and Air Force maintenance centers (www.realization.com), some hospitals within the U.K.'s NHS (www.healthcaretoc.com), Japan's Ministry of Land, Infrastructure and Transportation (MLIT) (www.afinitus.com/japanpw.html), and the TOC for Schools (www.tocforschools.com).

⁷Barnard's four methods (Barnard, 2003) for breaking any conflict, focuses stakeholders mainly on identifying and challenging the conflict assumptions (rather than the traditional TOC approach that gets stakeholders to challenge all the logical assumptions related to a conflict). The conflict assumptions (Why D jeopardize C?, Why D' jeopardize B?, Why D is in conflict with D' and Why there is not another to satisfy B and C?) are the real reason why the conflict exists, and therefore, if challenged directly provide an efficient and effective way to break any conflict.

8 Pilot Workshops #3 and #4 were facilitated by Professor Antoine van Gelder and Barry Urban who also have provided the ongoing support to the four cities after the kickoff workshops, with support of the waste experts Andrea Schultheis and Dieter Steinbach.

⁹Samples of the simulation games frequently used in these sessions can be downloaded from www.goldrattresearchlabs.com.

[10](#) Ray Immelman's book, *Great Boss, Dead Boss*, explains the importance of tribal behavior in organizations.

[11](#) Large organizations such as Boeing and ABB have shown that it is possible to do a “bottom-up” implementation of TOC. Generally, these do take longer and need a committed and growing group of internal TOC experts and practitioners, supported by external TOC experts, and an effective communication mechanism to share internal results and lessons learned with the rest of the organization.

¹PIMS is the acronym for Profit Impact of Market Strategies, a commercial database, originally developed for General Electric. The variables gathered included such items as market position, product quality, and capital investment. Studies were then conducted by many different researchers to determine the variables' relationships to such outcomes as profitability and market share. Eventually, the PIMS database included data on several thousand businesses.

²For more information on the exercise of power in the strategy-making process, see for example, Bower and Doz (1979), Brandenburger and Nalebuff (1995), Guth and Macmillan (1986), Hamel, Doz, and Prahalad (1989), MacMillan (1978), Pettigrew (1977), and Sarrazin (1977-1978).

³The *TOCICO Dictionary* (Sullivan et al., 2007, 43-44) defines *Strategy and Tactic tree (S&T)* as—A logic diagram that includes all the entities and their relationships that are necessary and sufficient to achieve an organization's goal. The purpose of the S&T tree is to surface and eliminate conflicts that are manifested through the misalignment of activities with organizational goals and objectives.

⁴Major measures for the Sales function relate to the amount of sales made; many times sales are made based on promises of short lead time. Production, on the other hand, is measured by efficiencies of resources, which translates into proper scheduling and large batches. Still another view is Finance which strives for low inventory levels which translates into few items stocked and little inventory of each item. The TOC logistics solution provides a win-win-win solution to these conflicts. DBR, for example, provides high utilization of the key resource, the constraint (satisfying production), rapid response to sales needs (short lead times and dependable service) and little investment in inventory (due to the responsiveness of DBR and the make-to-availability stocking option. The customer is the ultimate winner with this strategy.

⁵See the following Web site for examples of some generic S&T trees: <https://www.toc-goldratt.com/store/product.php?productid=111121>.

¹Scientific research shows that approximately 90 percent of business process re-engineering efforts fail to achieve tangible results (A.T. Kearney M&A study referenced in Bruner, 2004.). Fifty percent of all mergers and acquisitions fail outright, while most fail to realize expected synergies (McDonald, Coulthard, and de Lange, 2005). Fewer than 50 percent of companies undergoing restructuring or downsizing realize lower costs or higher productivity (Schneier, Shaw, and Beatty, 1992).

[2](#)See Section VI for a discussion of the TOC Thinking Processes.

³The TOCICO Dictionary (Sullivan et al., 2007, 43) defines strategy and tactic tree (S&T)—A logic diagram that includes all the entities and their relationships that are necessary and sufficient to achieve an organization's goal. The purpose of the S&T tree is to surface and eliminate conflicts that are manifested through the misalignment of activities with organizational goals and objectives.

Usage: Organizational strategy specifies the direction of the activities that purport to address longer range problems and issues. Tactics are the specific activities needed to achieve the strategic objective involved in implementing organizational strategies. Since strategy and tactics exist and must be synchronized within various organizational levels, this logic tree translates high level strategy down to the level of day-to-day operations. (© TOCICO 2007, used by permission, all rights reserved.)

⁴Some tables are mentioned in the text that are not shown in chapter tables. Use the Harmony to view the full S&T tree structure and tables. For example, this S&T tree is four levels (numbered in upper left corner) and 29 tables. Hence only a few tables are shown here for illustration. The Harmony viewer can be downloaded at: <http://www.goldrattresearchlabs.com/?q=node/2>

⁵**simplified drum-buffer-rope** (S-DBR)—The process of managing operations based upon a shipping buffer, a market drum, and secondary attention to any capacity constrained resources (Sullivan et al. 2008, 43).

⁶See Goal Systems International Website. Books and papers:
Simplified Drum-Buffer-Rope. 2000.
<http://www.goalsys.com/books/documents/S-DBRPaper.pdf>
for an excellent description of the concept by Schragenheim
and Dettmer.

[Z](http://www.furnituretoday.com/article/25111-California_s_Orman_Grubb_closing_doors.php)For the full story, see Furniture Industry News Web site, Furniture Today archives, Issue 32, April 12, 2004.
http://www.furnituretoday.com/article/25111-California_s_Orman_Grubb_closing_doors.php

[8](#)In such a chapter of a book, it is impossible to anticipate, let alone answer all possible questions related to an S&T. However, if you have questions, you are welcome to email them to gerryikendall@cs.com. I will do my best to provide meaningful answers to clear, concise questions.

⁹Typically, the only projects that are executed without the intention of bringing independent value are mandatory projects (e.g., those dictated by government regulations). Often, organizations split a single initiative into multiple projects, where some are prerequisites and do not bring value until the follow-on projects are completed. In the author's opinion, this is a grave mistake. From observation, often the prerequisite projects are completed with huge, sometimes indefinite, time lags before executing the follow-on projects that deliver to the bottom line.

10The TOC pull replenishment system also provides automatic adjustments, up or down, to inventory buffers at each link in the chain. When the actual inventory level at a location appears too frequently in the red zone, the target inventory level is automatically increased by one third. Similarly, when trends are down, the inventory will be in the green zone for most or all of the replenishment period. Then inventory targets are adjusted down by one third.

[11](#)The implication of having less inventory at the retail level is counterintuitive for manufacturers and distributors. In the TOC distribution paradigm, retail sales grow for all the reasons explained in this section. Therefore, manufacturers and distributors also gain greatly from increased sales and customer satisfaction and loyalty.

¹²Please read [Chapter 36](#), this volume.

¹The military context is the basis for this taxonomy, as reflected in [Table 19-2](#). In military applications, operations are large-scale coordinated events (often multi-service). Tactics are normally employed by smaller, discrete units.

²Many people and organizations make no concerted effort whatsoever to observe what's going on around them and put such observations into any kind of context relevant to themselves. As Winston Churchill once observed, "Man will occasionally stumble over the truth, but most of the time he will pick himself up and continue on" (Winston Churchill, <http://quotationsbook.com/quote/19633/>).

³It was nearly two decades before Boyd himself actually identified, analyzed, and articulated the OODA process he was unquestionably practicing it in the 1950s. But he was doing it all the same.

⁴Different people refer to the methodology created by Goldratt variously as *thinking process* or *thinking processes*. For the past eight years, I have inserted the word “logical” when I refer to it and used the singular form in order to more simply convey what the method involves to audiences having little or no prior exposure to TOC. The simplified, more streamlined version of the thinking process that I teach now—what amounts to a third generation—differs enough from Goldratt’s initial conception that I believe it warrants a modified name. The essential concept of logic trees, though, is still the brainchild of Goldratt.

[5](#) Current Reality Tree, Evaporating Cloud, Future Reality Tree, Prerequisite Tree, and Transition Tree.

[6](#)The use of the IO Map is not limited to strategy development alone. As it happens, its use as the first step in the LTP for *any* purpose is highly recommended. See Dettmer (2007) for a more detailed explanation.

Note that depending on environmental conditions, “maximum profitability” might actually be numerically negative. Nevertheless, it would be the smallest negative number possible to achieve.

⁸A cybernetic system is one that is affected by environmental shifts but has the means through feedback control to continue to meet system objectives. Additionally, a cybernetic system's objectives are not rigidly fixed but are adaptable to changing conditions and responsive to new understanding. Cybernetic systems gain from experience and thus exhibit learning (Athey, 1982).

⁹It's highly desirable to capture baseline figures, statistics, and other data in the first iteration of the *observe* step to facilitate effective detection of change in the second iteration of observation. Too often, this is neglected in actual practice.

[10](#) Other chapters in this Handbook provide guidance on constructing CRTs. *The Logical Thinking Process* (Dettmer, 2007) provides step-by-step explanation and instructions not found elsewhere specifically for integrating the IO Map with the CRT.

¹There are two distinct motives for initiating a change: (1) there is a problem in the current situation or (2) there is an opportunity we would like to seize or a vision we would like to pursue. In the latter, we ask Wary Will to climb up the cliff not because there is an alligator behind him but because there is a treasure up there. This situation requires a very different buy-in process that is outside the scope of this chapter. If we try to use the Layers of Resistance here, we might very well get stuck at Layer 0—the other party will insist there is nothing wrong in the current situation, which is actually correct.

²According to their suggestions, the more familiar people become with the change, the more control they have over decisions and actions, and the more they invest their time, ideas, and resources, the stronger their sense of ownership becomes.

¹In our company, we refer to sales opportunities as “projects.”

²I highly recommend each and every one of you to read this article. This paper has since been published: Goldratt, E. M. 2009. "Standing on the Shoulders of Giants". *The Manufacturer*. June. accessed Feb. 4, 2010 at http://www.themanufacturer.com/uk/content/9280/Standing_on_the_shoulders_of_giants.

³Used with permission by E. M. Goldratt (2009). © E. M. Goldratt, all rights reserved.

⁴There are many fun and insightful exercises that demonstrate the damage of bad multitasking. In our company, we particularly like to use the “paper-tearing” game.

⁵The way to read a Current Reality Tree is to start from the bottom of the tree upwards. You read “if [statement at the bottom of the arrow] then [statement at the top of the arrow].” If several arrows are tied together with an ellipse, then all the statements tied together at the bottom of the arrows should be read using “if [first statement] and if [second statement], then [statement at the top of the arrow].”

[6](#)To better understand this concept relating to production, please read “Standing On The Shoulders Of Giants,” talk to anyone in planning in case S-DBR is implemented in your company, or read Step 4:11 in a Strategy & Tactics Tree for Make-to-Order companies.

[Z](#)As of today, the only place Goldratt refers to this curve is in his latest Project Management Webcast series. As one would read in “Standing on the Shoulders of Giants,” the representative curve in most production environments is an inverse curve, referred to as the “U curve.” The following explanation is a paraphrase of Goldratt’s explanation of the U curve in “Standing on the Shoulders of Giants.”

$\frac{8}{250 \times (1 - 0.11)}$ is approximately 110 opportunities. Following the formula would have guided us to cut 45 percent of the projects. In reality, we cut 50 percent of the projects, which is practically the same.

⁹We intend to use the green, yellow, red status indicators, not as part of the POOGI, but as a daily management tool to identify delays early on and focus management attention before it accumulates to a significant delay.

10 While we claim low variability is an indication to a flaw in the process affecting the performance of all salespeople, we are not claiming that high variability necessarily indicates that the source of disruption relates to the individual performance of a salesperson. High variability in performance could be a result of different processes applied by salespeople, or that some of the salespeople don't follow a flawed process. When one experiences high variability in performance, both possible types of disruptions to flow should be explored.

¹[Send an email to refund@MafiaOffers.com](mailto:refund@MafiaOffers.com) to request your refund. You will be refunded the e-book price of this chapter.

²It is not a strong offer because it can be copied easily. However, it is unusual for such an offer to appear in a book and I sincerely believe that if you spend the time with this material, you can have a positive impact on your business.

³Originally a white paper titled Less is More, revised and now published in [Chapter 21](#) of this Handbook. Copyright © 2010 by Dr. Lisa Lang.

⁴See [Chapter 21](#) for an updated revision to this white paper.

⁵Firefighting mode is when you are consumed with emergencies and last minute priorities instead of planned, rational progression and improvement.

[6](#)Jayne Smith had similar findings and published them in her book, *Creating Competitive Advantage*.

[I](#)SCAs, USPs, and CVPs are often no more than tag lines.

8 You may realize some benefit to applying sales process management or funnel management even when you have a market constraint, but you will experience much bigger benefits with the combination of a Mafia Offer and funnel management. Funnel management (or sales process management) is typically handled by TOC practitioners by applying *Drum-Buffer-Rope* (DBR) to the sales process. See [Chapter 21](#) or <http://www.SalesVelocitySystem.com> where these concepts are applied.

⁹We can typically ask a few questions, depending on the type of operation, and get some idea of the improvements that are possible. However, if you do not have this experience, [Fig. 22-1](#) can serve as a guide.

[10](#) I am not saying we would start quoting a 2-week lead time, but that our internal cycle time would go from about 6 weeks to about 2 weeks and we will capitalize on our shorter internal lead time.

[11](#)Please refer to the logistics chapters (7 through 12) of this Handbook for further guidance on establishing operational improvements.

[12](#)For our analyses we use the rigorous cause-and-effect logic of the TOC thinking processes.

13Our version of S-DBR for custom job shops is called Velocity Scheduling System and you can find more information at www.VelocitySchedulingSystem.com

[14](#)This section draws heavily from the concepts in *Neuromarketing: Understanding the “Buy Buttons” in Your Customer’s Brain* by Patrick Renvoise’ and Christophe Morin, but what is new is the combining of neuromarketing concepts with TOC concepts.

[15](#)The solution for sales method for presenting offers was developed by Rami Goldratt and first presented at the 2003 TOC Upgrade Workshop in Cambridge, England.

[16](#) See excerpt from Stephan Shapiro's August 7, 2007 newsletter.

[17](#) Excerpt from Stephen Shapiro's August 7, 2007 newsletter.

18 Get a TOC Tip of the day at www.twitter.com/TOCExpert

¹⁹See www.MafiaOffers.com
www.MafiaOfferBootCamp.com

and

²⁰See www.MafiaOffers.com
www.MafiaOfferBootCamp.com

and

²¹See [Chapters 18](#), [25](#), and [34](#) on Strategy and Tactics trees.

[22](#)Goldratt Consulting, June 26, 2006, Vendor Management Inventory S&T.

²³See [Chapter 9](#) for more information on S-DBR.

²⁴See www.VelocitySchedulingSystem.com based on S-DBR for highly custom job shops.

[25](#)The TOC Replenishment Solution is also called Demand-Pull and more information can be found in [Chapter 11](#).

[26](#)Goldratt Consulting, November 2008, Reliable Rapid Replenishment S&T.

²⁷See [Chapter 9](#) for more information on S-DBR.

²⁸See www.VelocitySchedulingSystem.com based on S-DBR for highly custom job shops.

[29](#) Throughput Per Shelf (TPS) is a measure of return on shelf space.

[30](#) Goldratt Consulting, November 12, 2007, Consumer Goods S&T.

³¹See [Chapter 9](#) for more information on S-DBR.

³²See www.VelocitySchedulingSystem.com based on S-DBR for highly custom job shops.

[33](#)The TOC Replenishment Solution is also called Demand-Pull and more information can be found in [Chapter 11](#).

[34](#)Goldratt Consulting, June 21, 2007, Projects S&T.

[35](#) More information on the TOC CCPM solution can be found in Section II.

³⁶See www.ProjectVelocitySystem.com based on CCPM for service and project based companies.

[37](#)Goldratt Consulting, May 25, 2006, Pay Per Click S&T.

³⁸See [Chapter 9](#) for more information on S-DBR.

³⁹See www.VelocitySchedulingSystem.com based on S-DBR for highly custom job shops.

⁴⁰See www.ProjectVelocitySystem.com based on CCPM for service and project based companies.

[41](#) More information on the TOC CCPM solution can be found in Section II.

⁴²See [Chapter 13](#) for more on totally variable costs. Typically, they would be the raw material costs of the equipment.

⁴³See [Chapter 9](#) for more information on S-DBR.

⁴⁴See www.VelocitySchedulingSystem.com based on S-DBR for highly custom job shops.

[45](#)The TOC Replenishment Solution is also called Demand-Pull and more information can be found in [Chapter 11](#).

⁴⁶See www.ProjectVelocitySystem.com based on CCPM for service and project based companies.

[47](#) More information on the TOC CCPM solution can be found in [Chapters 3, 4, and 5](#).

⁴⁸See www.MafiaOffers.com
www.MafiaOfferBootCamp.com.

and

⁴⁹Based on experience from the over 70 companies that have completed a Mafia Offer Boot Camp. See www.MafiaOffers.com and www.MafiaOfferBootCamp.com.

¹Attributed variously to Henri Poincaré, James Maxwell, and Kurt Lewin.

²Cox, Blackstone, and Schleier (2003, 47-61) developed their Business System Model to describe the system prior to implementing TOC; Lockamy and Cox (1994, 11) address “What is the Goal?” and “How is it measured?”

³Why change and establishing a POOGI are addressed by Barnard in [Chapter 15](#); sustainability of change is addressed by Newbold in [Chapter 5](#).

⁴Note that for the analysis by Kim et al. (2008), work contained in books was excluded due to the inherent difficulty of identifying individual examples mentioned in books. However, books comprise a major component of the TOC literature, being almost equal in number to the number of papers surveyed in Kim et al. Up until 1998, there were approximately 28 books on TOC; in the last 10 years, about 70 more books have been published, including a dozen on TP and a similar number of educational workbooks on TP.

[5](#)The original early 1990s method proposed to develop a CRT used “Why?” and “Because” logic with validation, and is still in use.

[6http://tocinternational.com/pdf/Oxford%20Radcliffe%20Hospital%20story.pdf](http://tocinternational.com/pdf/Oxford%20Radcliffe%20Hospital%20story.pdf)

[Z](http://www.goldrattresearchlabs.com/)To review the S&T trees that currently are in the non-peer-reviewed public domain see: <http://www.goldrattresearchlabs.com/>

⁸A basic set of TOC logic tools (EC, NBR, and PRT) has been taught in primary, secondary, and high schools around the world for over a decade. See [Chapter 26](#) and www.tocforeducation.com.

1The Cloud is used in the full TP work to describe the inherent conflict reflected in the core problem that is identified in the *Current Reality Tree* (CRT).

²*The TOCICO Dictionary* (Sullivan, Reid, and Cartier, 2007, 21-22) defines the *Evaporating Cloud (EC)* as " (a) necessity-based logic diagram that describes and helps resolve conflicts in a "win-win" manner. It has two primary uses, first as a structured method to facilitate the description and resolution of a conflict, and second, as an integral part of the Three-Cloud approach to creating a Core Conflict Cloud which then forms the base of a current reality tree." (© TOCICO 2007, used by permission, all rights reserved.)

³In some instances, particularly Clouds involving more than one person, the I/we should be replaced with the part's name if that person or function must meet that need or complete that action.

⁴Chronic Conflict Clouds are far more difficult to resolve. The *TOCICO Dictionary* (Sullivan et al., 2007, 11) defines *chronic conflict* as “(a) contentious situation that has continued to exist for a prolonged period of time. Opposing sides have been justifying their perspective through selective requirements and prerequisites for so long that both sides become entrenched in their own beliefs to the point that neither side can see how to break the conflict without suffering a significant loss.” (© TOCICO 2007, used by permission, all rights reserved.)

⁵The *TOCICO Dictionary* (Sullivan et. al., 2007, 27-28) defines *injection* with respect to the EC as: “2. A state or condition that invalidates one or more assumptions underlying the relationships between the objective and requirements, or between requirements and prerequisites, or between the two prerequisites of an Evaporating Cloud.” (© TOCICO 2007, used by permission, all rights reserved.)

[6](#)Used with permission by Eliyahu M. Goldratt. © Eliyahu M. Goldratt.

⁷See, for example, [Chapter 20](#), this volume, Dr. Efrat Goldratt's chapter on layers of resistance and buy-in.

⁸In [Chapter 2](#) of the book *It's Not Luck* by Eli Goldratt (1994), there is an example of a day-to-day conflict on a personal issue. It is a disagreement between Alex Rogo and Sharon, his teenage daughter.

⁹Please remember that this is a daily problem and not a detailed analysis of a deeper problem. A more comprehensive work may find that the procedure is a part of the core problem and in such a case, the solution may include a major change or even the removal of procedure.

¹⁰The *TOCICO Dictionary* (Sullivan et al., 2007, 50) defines an *undesirable effect (UDE)* as “(a) negative aspect of the current reality defined in relation to the organizational or system’s goal or its necessary conditions. UDEs are believed to be a visible symptom of a deeper, underlying root cause, core problem, or core conflict.” (© TOCICO 2007, used by permission, all rights reserved.)

¹¹See [Chapter 22](#) for details on constructing Mafia Offers.

[12](#)The Three-Cloud approach is also used in examining UDEs across major functions of an enterprise. However, these analyses do not fall into the category of day-to-day management and therefore are not discussed in this chapter.

¹³The *TOCICO Dictionary* (Sullivan et al., 2007, 14) defines “core conflict—The systemic conflict that causes the vast majority of the undesirable effects in the current reality of the system being studied. The core conflict is often generic in nature and can be derived by generalizing the various conflicts that underlie the undesirable effects that persist in the system.” (© TOCICO 2007, used by permission, all rights reserved.)

[14](#)Please note that the core conflict within the subject matter can be identified through the same process of consolidating all of the identified UDEs as reflected in the U-Shape discussed later. Please note also that any analysis is sensitive to the list of UDEs that is picked. Therefore, it is important to check that the UDEs have severe impact on the performance of the area under study.

¹⁵Dr. Eli Goldratt has described the basic concepts/assumptions of TOC in numerous presentations. Very little has been written about these assumptions. In his presentation called “Necessary & Sufficient” on the second part (on CD-2), he presents the first two of the concepts under the heading: “The basic assumptions of TOC—A look into reality based on the common sense approach of the hard sciences.” The three basic assumptions also are described in detail in Goldratt’s (1990) book *What Is This Thing Called Theory of Constraints and How Should It Be Implemented?* The same basic assumptions are covered in *The Choice* (Goldratt, 2009). They are recorded on an extended list of his assumptions on pages 157–158 of the book.

¹⁶A nice example of such a case appears in [Chapter 8](#) of *It's Not Luck* (Goldratt, 1994).

[17](#)If you have many IOs, you should consider grouping them and establishing group topics and complete the exercise for each group. Then, sequence the groups.

¹As I write, I am imagining more than one occasion in which I thought a TV, a computer, or other electronic gadget was not working, and “the fix” came in the form of my husband calling from another room with just a small tinge of sarcasm in his voice, “Honey, are you sure it’s plugged in?”

²The detailed process is called the *Categories of Legitimate Reservation* (CLR).

³The And Connector was originally named (and is still often called) a “banana” due to the shape that is formed when writing trees by hand to illustrate the “logical and” nature of the causes and effects.

⁴Entities 1, 3, 5, 6, and 9 are causes only. Entities 11 and 12 are effects only. Entities 2, 4, 7, 8, and 10 are both causes and effects. Entities 1 and 10 are a cause for Entity 2; Entities 2 and 3 are a cause for Entity 4; Entity 4 is a cause for Entity 7; Entities 6 and 7 are a cause for Entity 10; Entity 10 is a cause of Entity 2 (the loop); Entity 7 is a cause for Entity 11; Entities 4 and 5 are a cause for Entity 8; Entity 8 is a cause for Entity 12; and Entity 9 is another cause for Entity 12.

[5](#) More on the term *undesirable effect* (UDE) can be found in the section of this chapter that describes the Current Reality Tree.

⁶When a *Strategy and Tactic Tree* (S&T) is used, it often replaces the PRT as the mechanism to address the obstacles to implementation of a solution that contains many injections and requires the synchronization of multiple stakeholder groups.

⁷If you are interested in learning the history behind the quite different name of this TP, see [Chapter 9](#) of the book *Thinking for a Change: Putting the TOC Thinking Processes to Use* (Scheinkopf, 1999).

8 A source of delay is the answer to the question, “What is the order waiting for?”

⁹In Appendix C, I provide a copy of the POOGI step in the standard Strategy & Tactic Tree that MTO manufacturers implement. See

<http://www.mhprofessional.com/TOCHandbook>

10The TOC Priority Management (using BM) approach classifies orders as green, yellow, red, or black according to the degree to which the order has consumed its buffer (safety time). Green orders have consumed the least amount of buffer (and are thus not at risk of being late), black orders have consumed their entire buffer (and are thus already late).

11Note that you are looking for assumptions that are being made in the given situation. Therefore, it may be helpful for you to modify the statements listed here to include the words “it is believed” or “it is thought.” For example, “In order for A, it is believed that we must B, because”

[12](#) These four rules summarize the CLR, which are described in detail in Appendix B at the end of the chapter.

[13](#)In the United States, urgent care centers are non-hospital medical clinics where people who do not have a primary care physician or whose physician is not available can go.

[14](#)Just in case you are wondering, the young woman subsequently had the tumor removed, was diagnosed with aggressive brain cancer (glioblastoma multiforme), and continues to outlive the statistics that are otherwise translated to be a death sentence.

¹⁵John Covington (in [Chapter 37](#)) discusses this application.

16By “real UDE,” I do not mean “real causes.” I do mean the entities that are undesirable on their own merit. In the next step, you will identify the core problem—the cause that is responsible for the existence

[17](#)The PRT goes by a number of names and procedures. For instance, it is described as an Ambitious Target Tree in [Chapter 26](#), and its derivative, *Intermediate Objective Map* (IO Map), is presented in [Chapter 24](#).

18 If you are going to create an S&T tree, then this third step is not necessary.

[19](#)Over the years, several formats for constructing TRTs have been developed, taught, and used. The version that I am including in this chapter is different—and in my view more effective relative to its purpose—than the version I presented in *Thinking for a Change*. As always, though, if you keep the view of the scientist and your objective in mind, you have the opportunity to develop an approach that works for you.

[20](#)An objective can be an injection from an FRT, an IO from a PRT, or any other objective that you would like to achieve that did not arise from a full TP analysis.

[21](#)Increasing value: stability on the green curve, performance on the red curve. See [Chapter 34, Fig. 34-1](#).

[22](#) Step 1, Viable Vision, used with permission of Goldratt Consulting.

[23](#)The lowest level steps in an S&T do not contain a Sufficiency Assumption.

²⁴In [Chapter 34](#), Lisa Ferguson provides several detailed examples of S&Ts.

²⁵© E. M. Goldratt (2008) used by permission, all rights reserved. Source: Goldratt Research Labs at: <http://goldrattresearchlabs.com/?q=node/2>

²⁶Thanks to Stewart Witt for his contribution, and to Goldratt Consulting and Revital Cohen for their permission to use this TRT. The rest of the TRT is in Appendix E at the McGraw-Hill Website: <http://www.mhprofessional.com/TOCHandbook>

[27](#) Also called Layers of Resistance.

[28](#)From Cox et al., 2003, pp. 83–88. Used with permission. © Cox, Blackstone and Schleier. All rights reserved. Appendixes A and C to G are on the website

¹Two such examples are Anaya and Pamanes, “Violence in the Home” at <http://www.tocforeducation.com/cloud-b/cb23.html> and de Gaza and Rodriquez, “Enabling Juvenile Offenders to Set Goals” at <http://www.tocforeducation.com/att-b/attb09.html>

[2](#)“**Mainstreaming** in the context of education is a term that refers to the practice of educating students with special needs in regular classes during specific time periods.”
Wikipedia. Copyright © 2010 by Kathy Suerken.

³In a presentation by a local third grade teacher (Anonymous, 1994), she revealed how TOC made her realize that the materials and techniques she had been using to teach cause-and-effect logic were fundamentally flawed. In so doing, she quipped, “What should I do now? Write all my students letters of apologies?”

⁴Brazil, Colombia, Costa Rica, Ecuador, Israel, South Korea, Mexico, Malaysia, The Netherlands, Philippines, Poland, Peru, Russia, Serbia, Singapore, Republic of South Africa, Taiwan, Trinidad & Tobago, United Kingdom, United States, Venezuela.

[5](#) Keynote speech to 1st TOCfE International Conference, Los Angeles, CA, August 1997.

⁶The TOCfE *Thinking Across the Curriculum Workbook Series* defines the Cloud as a logical thinking diagram that defines and analyzes a problem through different points of view in a way that eliminates the conflict without compromising important needs; the Logic Branch as a logical thinking diagram that describes through cause-effect relationships how an entry point leads to outcomes; and the Ambitious Target Tree as a logical diagram to construct feasible strategic and tactical plans to attain an ambitious target by analyzing obstacles and developing specific, sufficient, and sequenced steps that turn stumbling blocks into stepping stones. *Thinking Across the Curriculum Series*, Suerken, ©TOCICO for Education, Inc, 2009. See also *The TOC Dictionary* (Sullivan et al., 2007) at http://www.tocico.org/resource/resmgr/files-public/toc-ico_dictionary_first_edit.pdf

[Z](http://www.tocforeducation.com/cloud-b/cb7.html)Monterrey, Mexico, August 2000. Full case study entitled “TOC Mediation to Stop Name Calling” is available at <http://www.tocforeducation.com/cloud-b/cb7.html>

⁸In TOC, an *assumption* is a statement, condition, or belief about why a logical relationship exists between entities.

9“An Alternative to Hazing,”
<http://www.tocforeducation.com/cloud-b/cb2.html>. Roby continues to use the Cloud with students and reports that a similar hazing situation occurred on the first day of school in the fall of 2009. He took particular note of the insights gained through the use of the Cloud by the senior girls involved in the incident. When verbalizing their needs, they discovered that their need to be respected was actually being jeopardized by actions to haze others.

[10](#)Speech NCA Conference, Chicago, IL, April 1997. The North Central Association Commission on Accreditation and School Improvement accredits schools in 19 states.

[11](#)“AGOAL Academy” Presentation of Ana Maria Conde, 8th TOCfE International Conference, Seattle, WA, August 2005.

¹²Trapnell's action research on the use of the TOC tools at Alderman Pounder Nursery School has been published in *Child Education* (August 1998), *Primary Leadership Paper* (January 2003), and *Teaching Expertise* (Winter 2004).

[13](#)The Free Online Dictionary defines *inference* as “the act or process of deriving a logical consequence or conclusion from existing premises” (www.thefreedictionary.com/inference).

[14](http://www.tocforeducation.com/cloud-c/cc01.html)<http://www.tocforeducation.com/cloud-c/cc01.html>

[15](#)Glatter, 2003 “Reading Comprehension Through TOC,”
Presentation 7th International Conference, Ft. Walton Beach,
FL (October).

[16](#)“Ft. Walton Beach, FL, October 2007.

[17](http://www.tocforeducation.com/references.html)<http://www.tocforeducation.com/references.html>

[18](#) Presentation of Marilyn Garcia to the Maryland Home Education Association, November 2006.

¹⁹Combined from “The Case of the Disruptive Student,” <http://www.tocforeducation.com/branch-b/bb01.html> and Small’s Presentation at 7th TOCfE International Conference, Ft Walton Beach, FL, May 2003.

[20](http://www.tocforeducation.com/branch-b/bb02.html)“I Have Had No Further Problem with Tardiness,”
<http://www.tocforeducation.com/branch-b/bb02.html>

[21](#)“TOC in Counseling: Taking Responsibility for Learning, A Classroom Behavior Intervention,” Marcia Hutchinson:
<http://www.tocforeducation.com/att-b/math3.html> and
<http://www.tocforeducation.com/att-b/math4.html>

[22](http://www.tocforeducation.com/branch-b/bb03.html)<http://www.tocforeducation.com/branch-b/bb03.html>

[23](#)Languages spoken include Malay, Tamil, Chinese, and English.

[24](http://www.tocforeducation.com/cloud-b/cbl2.html)<http://www.tocforeducation.com/cloud-b/cbl2.html>

[25](#) TOCfE in Malaysia, 2004 Presentation at the 7th TOCfE International Conference, Ft. Walton Beach, FL, May; 2006 “100 Children ×100 Days × 100 Clouds” Presentation at the 9th TOCfE International Conference, Leon, Mexico, Sept.; 2005 *Thinking Smart: You Are How You Think*, Selangor Malaysia: Pelanduk Publications.

[26](http://www.tocforeducation.com/att-b/attb02.html)“Changing the Mindsets of Groups of Disruptive Students,”
<http://www.tocforeducation.com/att-b/attb02.html>

[27](#)Dr. Sirias has co-authored a TOC book for teenage children entitled *SUCCESS: An Adventure* and is currently developing a new workshop that incorporates TOC in mathematics delivery.

[28](#) Florida Comprehensive Assessment Test.

[29](#) Presentation to 7th TOCfE International Conference, Ft. Walton Beach, FL, May 2003.

[30](#)“TOC and the Children of Romani Populations,”
Presentation at the 11th TOCfE International Conference,
Warsaw, Poland, October 2008.

[31](#)This application, developed by Christina Cheng, is described in [Chapter 27](#).

[32](#)The methods used to evaluate level of social maladjustment were the Rosenzweig Picture Frustration Test and Student Behavior Test of Barbara Markowska. In order to evaluate the level of solving conflicts, predicting both positive and negative consequences of someone's behavior and planning small undertakings, a set of TOC tasks were created for the needs of the experiment. Conference Proceedings, 10th TOCfE International Conference, Ft Walton Beach, FL, October 2007.

[33](http://www.tocforeducation.com/researchlist.html)To view the full presentation:
<http://www.tocforeducation.com/researchlist.html>

[34](#)Conference Proceedings, 8th TOCfE International Conference, Seattle, WA, August 2005.

[35](http://www.tocforeducation.com/researchlist.html)To view “Curriculum Applications” presentation that includes summary of results: <http://www.tocforeducation.com/researchlist.html>

[36](#)The *Thinking for a Change* workbooks for children (in Hebrew) include: *Solving Every Day Conflicts* for ages 5-8, 8-12, and 10-15; *The Way of Achieving a Target* for ages 8-12 and 10-15; and *Think Before You Act* for ages 5-8 and 8-12. Glatter, along with Mira Grienberg and Rami Goldratt, has also written *Rainbow in the Cloud*, a workbook for teachers.

[37](http://www.tocforeducation.com/interactive.html)<http://www.tocforeducation.com/interactive.html>

[38](#)Avraham Y. Goldratt Institute. 1995. Management Skills Workshop Sessions 1-5, New Haven, CT: Avraham Y. Goldratt Institute.

[39](#) Some nomenclature was changed in order for terminology to be user friendly for children, and learning objectives have been added and tailored to meet the needs of educators. However, all adaptations in language and content of the original materials have been carefully undertaken and approved by the creator of TOC, Dr. Goldratt, in order to prevent distortions in the processes that could affect desired intended outcomes.

[40](#) Her thesis is entitled, “The Effectiveness of Integrating the Theory of Constraints for Education in the Teaching Learning Processing English I,” 126.

[41](#)MSCDN, Mazowieckie Samorządowe Centrum
Doskonalenia Nauczycieli, services 50,000 Polish teachers.

[42](#) Los Angeles School District, CA.

[43](#) Presentation to 3rd International TOCfE Conference Los Angeles, CA, August 1999. As advisor to the Officer of Intergroup Relations, LASD, CA, Denise co-authored “La Crème: Los Angeles Conflict Resolution Education Model for Educators.”

[44](http://www.tocforeducation.com/att-c/attc04.html)Linda Trapnell, “Learning How to Make Logical Plans”:
<http://www.tocforeducation.com/att-c/attc04.html>

[45](#)Dr. Cora Santiago and Lourdes Visaya, Presentation at the 6th TOCfE International Conference, Nottingham, England, July 2002.

[1](#)Source: **Mercer Quality of Living global city rankings 2009*, the latest Trends in International Mathematics and Science Study (TIMSS).

Copyright © 2010 by Christina Cheng.

²The chapter reports the actual experiences of the project and is not intended as a formal academic study.

³Participants were asked what obstacles were blocking them from achieving the ambitious target “To be successful in the workplace.”

⁴The Yellow Ribbon Project is spearheaded by the Community Action for the Rehabilitation of Ex-Offenders (CARE) Network, a group of major community and government organizations responsible for the rehabilitation of ex-offenders.

⁵Refers to the work of Edward Lorenz based on chaos theory whereby the flap of a butterfly's wings may contribute to a tornado in another part of the world by creating tiny changes in the atmosphere.

⁶As a funny aside, we were puzzled when used flipchart paper kept disappearing from the trash. Later we discovered some inmates had taken the examples to paste on their cell wall to show off their skills in-between classes to non-TOC cellmates!

¹Goldratt (Goldratt, Schragenheim, and Ptak, 2000) examines the use of technology in a business in his novel, *Necessary But Not Sufficient: A Theory of Constraints Business Novel*.

²The *Theory of Constraints International Certification Organization* (TOCICO) offers certifications in several areas including Supply Chain Logistics, Finance & Measures, Project Management, Thinking Process, and Business Strategy. Visit their Web site for further information.
<http://www.tocico.org>

³Literature reviews are provided for each of the TOC application areas at the beginning of each section in this handbook.

¹The figures in this chapter and some of the discussion are based upon material presented by the authors in their book *Release the Hostages*, published by North River Press in 2009. We wish to thank North River Press for their kind permission to present it here.

²*Customer Support Services* is known by several names including customer support, customer services, customer support services, technical support, or technical services. We will use the term customer services, or CS for short.

Copyright © 2010 by Alex Klarman and Richard Klapholz.

³The Thinking Processes are not taught in this chapter; it is assumed that the reader either is familiar with the processes or can read Section VI of the Handbook.

¹While not called Theory of Constraints at the time, many of the concepts are presented in Goldratt (1984).

²See Kendall (2004).

[3](#)For a discussion of Strategy and Tactic Trees, see Chapters 15, 18, 25 and 34.

Copyright © 2010 by Gary Wadhwa.

⁴An excellent discussion of Critical Chain in a multi-project environment is provided in Kendall and Rollins (2003).

⁵See Pascal Dennis (2007) and Sayer and Williams (2007) for descriptions and examples of the use of lean tools.

[6](#)See Mikel and Schroeder (2000) and Gygi, DeCarlo, Williams, and Covey (2005) for descriptions and examples on Six Sigma.

[Z](#)We run into dilemma and debate on what level of care must be provided to everyone. Where should society draw a line between mandatory health care verses voluntary individual choice of care? Should government provide health care for all or should we allow the free market to provide high quality, reliable care? This chapter does not get into a political debate. It does discuss the need to speed up the cure rate and provide reliable care.

[8](#)© E. M. Goldratt used by permission, all rights reserved.
(For a full development of this POOGI, see the section on
performance measurement.)

9Throughput takes into account variable costs of supplies for each service, the time utilization, number of patients served, and dollars paid by insurance companies.

[10](#)Please note, there is a strong assumption built into this argument. The private practices and hospitals can improve their Throughput significantly by the first two steps that the low value T/CU segment of population will become a significant source of profits. This is not different from the airline or hotel industries that try to fill the capacity by offering discounts through Priceline, Orbitz, etc. For-profit organizations have to discriminate and make rational decisions based upon TA in order to show bottom-line results.

[11](#)Example: For surgical procedures, the patients are more tolerant if the surgery prior to them is delayed. They have been fasting, taken time from work, and they are not interested in rushing the surgeon to perform faster, whereas during quick consulting or postsurgery follow-up time, any delay appears longer. The patients' expectations are to get in and out so that they can go on with their lives.

¹²The *TOCICO Dictionary* (Sullivan et al., 2007, 47) defines *Throughput (T)* as “The rate at which the system generates ‘goal units’. Because throughput is a rate, it is always expressed for a given time period such as per month, week, day or even minute. If the goal units are money, throughput will be an amount of money per time period. In that case throughput is calculated as revenues received minus totally variable costs divided by the chosen time period.

Illustration: Suppose a company produces only one product, and it sells for \$100 and has totally variable costs of \$35 per unit. If, in a week, the company produces 500 units but only sells 450, throughput would be \$29,250 per week $((100-35) \times 450)$. NOTE: Product produced but not sold does not generate throughput, it increases inventory.” (© TOCICO 2007, used by permission, all rights reserved.)

¹³The *TOCICO Dictionary* (Sullivan et al. 2007, 35) defines *Operating Expense (OE)* as “All the money the organization spends in generating “goal units”. Perspective: In the throughput-world paradigm of the theory of constraints, operating expenses include items such as salaries, rent, insurance, and other expenses that would be paid even if operations stopped for awhile. OE does not include expenses that vary directly with production/service volume, such as cost of raw material, commissions, etc. These expenses are considered to be totally variable costs, not OE.” (© TOCICO 2007, used by permission, all rights reserved.).

[14](#)The TP are discussed in detail in other chapters in this section. An Evaporating Cloud and the S&T related to the doctor's application are shown here to illustrate their application to the health care field.

[15](#) There is usually a lag time to see the results from actions taken after implementing logically valid initiatives. This delay causes the dynamics and dance between one side of the conflict and the other that results in negative feedback loops with associated undesired side effects.

[16](#)Some health providers may be upset by the focus on the goal of a for-profit health provider being to make more money now and in the future. They should be informed that any for-profit business must satisfy two necessary conditions to meet this money goal in the long term. These necessary conditions are that the provider must provide high quality service to the customer (patient in our health care environment) and to have satisfied employees (paid staff in our health care environment). Achieving the goal of making more money now and in the future is impossible unless these two necessary conditions are met. Hence, having already met these two necessary conditions is assumed and any future decision does not jeopardize this.

¹⁷TA is discussed in [Chapter 14](#) and also Corbett (1998).

¹⁸The *TOCICO Dictionary* (Sullivan et al., 2007, 29) defines *Investment (I)* as “All the money currently tied up in the system. As used in TOC, investment refers to the equipment, fixtures, buildings, etc. that the system owns as well as inventory in the forms of raw materials, work in progress, and finished goods.” (© TOCICO 2007, used by permission, all rights reserved.)

[19](#)The S&T replaces the Prerequisite Tree and the Transition Tree in developing detailed plans for an organization. It provides the basis for a detailed implementation plan to achieve the viable vision for any system. The health care S&T Tree is a blend of the Reliability, Rapid Response S&T and the Project Management S&T templates.

[20](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 24) defines *flying pig injection*—“A breakthrough solution or injection that initially seems impossible to implement.” (© APICS 2008, used by permission, all rights reserved).

[21](#) Definition is yet to be done for 4:31, 4:32, and 4:33.

¹Compromise (noun): something accepted rather than wanted. Compromise (verb): lessen the value of somebody or something. *Encarta Dictionary: English (North America)* July 2009.

²Throughout this chapter, the word *clinicians* is used to represent *all medical professionals who provide medical services to patients, including physicians, nurses, technicians, etc.*

³The *TOCICO Dictionary* (Sullivan et al., 2007, 21-22) defines *core problem* as “(a) fact, or conflict, or erroneous assumption that is the source of at least 70% of the undesirable effects in the current reality of the system being studied. Perspective: A core problem can have three manifestations either as 1. a fact, such as ‘efficiency is used as the prime measure in operations,’ or, 2. the conflict between *D* and *D'* in a core conflict cloud, such as ‘*D*. Use local efficiencies as a prime measure, and *D'*. Do not use local efficiencies as a prime measure,’ or, 3. an erroneous assumption responsible for the conflict, such as, ‘A resource standing idle is a major waste. ’” (© TOCICO 2007, used by permission, all rights reserved.)

⁴When one discusses an improvement initiative with employees, many times the employee will say: “That will never work here, we are different.” Pay close attention to the employee’s reasons it will fail; he or she is probably right. You may not be addressing the core problem, you may not understand what would block that initiative, etc.

⁵The *TOCICO Dictionary* (Sullivan et al., 2007, 11) defines *undesirable effect* as “A negative aspect of the current reality defined in relation to the organizational or system’s goal or its necessary conditions. UDEs are believed to be a visible symptom of a deeper, underlying root cause, core problem, or core conflict. Usage: Some characteristics of a well-articulated UDE include: 1. a complete statement about a single consequence which does not contain the following words/phrases: ‘and’, ‘because of’, or ‘as a result of ’; 2. an effect that is within management’s span of control; 3. something that exists in the reality of the organization precisely as stated; 4. something that is negative in its own right, without dependence on any other factor; 5. neither a presumed cause nor a presumed solution of the organization’s core conflict or its major dilemma. Most, if not all, UDEs should appear as entities within the current reality tree.” (© TOCICO 2007, used by permission, all rights reserved.)

[6](#) Evaporating Clouds are presented in Section VI on Thinking Processes, this volume.

[1](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 27-28) defines the *three-cloud approach* as “(a) relatively fast method of developing a current reality tree (CRT) wherein the developer identifies three seemingly independent undesirable effects (UDEs), creates an evaporating cloud (EC) for each, and synthesizes the three ECs into single generic cloud called the core conflict cloud (CCC).” (© TOCICO 2007, used by permission, all rights reserved.)

⁸The *TOCICO Dictionary* (Sullivan et al., 2007, 50) defines *change sequence* as “The three stages that must be completed in the successful management of change within a system. The change sequence answers the following three questions: 1. What to change? 2. To what to change? and, 3. How to cause the change?” (© TOCICO 2007, used by permission, all rights reserved.)

⁹The *TOCICO Dictionary* (Sullivan et al., 2007, 14) defines the *current reality tree* (CRT) as “(a) *thinking processes sufficiency-based logic diagram that facilitates answering the first question in the change sequence, namely, “what to change?”* The CRT is a diagram that illustrates the cause-effect relationships that exist between the core problem and the most, if not all, of the undesirable effects (UDEs).” (© TOCICO 2007, used by permission, all rights reserved.)

[10](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 8) defines *categories of legitimate reservation (CLR)* as “*The rules for scrutinizing the validity and logical soundness of thinking processes logic diagrams. Seven logical reservations are grouped into three levels. Level I: clarity reservation. Level II: causality existence and entity existence reservations. Level III: cause insufficiency, additional cause, predicted effect existence, and cause-effect reversal (tautology) reservations.*” (© TOCICO 2007, used by permission, all rights reserved.) The CLR are presented in [Chapter 25](#).

[11](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 14) defines *core conflict* as “(t) he systemic conflict that causes the vast majority of the undesirable effects in the current reality of the system being studied. The core conflict is often generic in nature and can be derived by generalizing the various conflicts that underlie the undesirable effects that persist in the system.” (© TOCICO 2007, used by permission, all rights reserved.)

[12](#)Such as nurses having to double chart—keeping both paper and computerized records of the same events. Another example might be an unacceptable time delay in the delivery of consumables from stores to the treatment areas, which results in staff having to beg or borrow supplies from other areas.

[13](#) How to use these thinking process tools is discussed in detail in Section VI on the Thinking Processes.

¹⁴The *TOCICO Dictionary* (Sullivan et al., 2007, 21) defines the *evaporating cloud (EC)* as “(a) necessity-based logic diagram that describes and helps resolve conflicts in a ‘win-win’ manner. It has two primary uses, first as a structured method to facilitate the description and resolution of a conflict, and second, as an integral part of the three cloud approach to creating a core conflict cloud which then forms the base of a current reality tree.” (© TOCICO 2007, used by permission, all rights reserved.)

¹⁵The *TOCICO Dictionary* (Sullivan et al., 2007, 34) defines the *negative branch (NBR)* as “(a)n adverse or undesirable side effect that can be caused by an injection and thereby compromise the positive effects from a proposed problem solution or injection.” (© TOCICO 2007, used by permission, all rights reserved.)

[16](http://www.tocforeducation.com/teach3.html)<http://www.tocforeducation.com/teach3.html>

[17](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 38) defines the *prerequisite tree (PRT)* as—“A necessitybased logic diagram that facilitates answering the third question in the change sequence, namely, how to cause the change? A PRT shows the relationship between the injections, intermediate objectives, or ambitious target, and the obstacles that block the implementation of the injections. A PRT includes the intermediate objectives required to overcome the obstacles, and shows the sequence in which they must be achieved for successful implementation.” (© TOCICO 2007, used by permission, all rights reserved.)

¹⁸The *TOCICO Dictionary* (Sullivan et al., 2007, 17) defines *desirable effect (DE)* as “(a) positive or beneficial outcome associated with an organization’s actual or future performance. DEs are often the opposite of an UDE.” (© TOCICO 2007, used by permission, all rights reserved.)

¹⁹The *TOCICO Dictionary* (Sullivan et al., 2007, 43) defines the *strategy and tactic tree (S&T)* as “(a) logic diagram that includes all the entities and their relationships that are necessary and sufficient to achieve an organization’s goal. The purpose of the S&T tree is to surface and eliminate conflicts that are manifested through the misalignment of activities with organizational goals and objectives.” (© TOCICO 2007, used by permission, all rights reserved.)

[20](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 15) defines *critical chain project management (CCPM)* as “The TOC solution for planning, scheduling, and managing performance in a project environment. It is applied in two very different environments; single project environments and multi project environments where resources are shared across several different projects concurrently.” (© TOCICO 2007, used by permission, all rights reserved.) See Chapters 3, 4, and 5, this volume.

²¹The *TOCICO Dictionary* (Sullivan et al., 2007, 17) defines the TOC *distribution/replenishment solution* as “A pull distribution method that involves setting stock buffer sizes and then monitoring and replenishing inventory within a supply chain based on the actual consumption of the end user, rather than a forecast. Each link in the supply chain holds the maximum expected demand within the average replenishment time, factored by the level of unreliability in replenishment time. Each link generally receives what was shipped or sold, though this amount is adjusted up or down when buffer management detects changes in the demand pattern.” (© TOCICO 2007, used by permission, all rights reserved.) See Chapters 11 and 12, this volume.

²²The *TOCICO Dictionary* (Sullivan et al., 2007, 18) defines *drum-buffer-rope (DBR)* as “*The TOC method for scheduling and managing operations.*” (© TOCICO 2007, used by permission, all rights reserved.) See Chapters 8, 9, and 10, this volume.

²³The *TOCICO Dictionary* (Sullivan et al., 2007, 7) defines *buffer management (BM)* as “A feedback mechanism used during the execution phase of operations, distribution, and project management that provides a means to prioritize work, to know when to expedite, to identify where protective capacity is insufficient, and to resize buffers when needed.” See [Chapter 8](#). (© TOCICO 2007, used by permission, all rights reserved.) See [Chapter 8](#), this volume.

²⁴The *TOCICO Dictionary* (Sullivan et al., 2007, 46) defines the *thinking processes (TP)* as “A set of logic tools that can be used independently or in combination to address the three questions in the change sequence, namely, 1. What to change? 2. To what to change? and, 3. How to cause the change? The TP tools are: evaporating cloud, current reality tree, future reality tree, negative branch reservation, prerequisite tree, and transition tree.” See Chapters 34 and 35. (© TOCICO 2007, used by permission, all rights reserved.) See Chapters 34 and 35, this volume.

[25](#)A recent CRT analysis incorporated the UDEs given by 65 staff. If these people were to schedule 30-minute interviews to voice their concerns with each other, it would have consumed over 2000 people hours and they would not have discovered the underlying core problem of their facility.

¹This definition is consistent with Eliyahu M. Goldratt's definition of *complexity*, "*Complexity is a result of the number of interactive constraints—constraints that impact each other.*" (Goldratt 1987, 1988, 1989, 1990, [Chapter 5](#), "How Complex Are Our Systems," 1.)

²In CCPM with single projects, all conflict is removed from the plan by moving tasks earlier in time, which extends the project but gives an aggressive yet feasible plan with high probability of completion within the project buffer. In sequencing multiple projects, there are frequent resource conflicts between projects. Trying to resolve the conflict between projects would force the projects to move conflicting tasks backward in time, which lengthens the individual projects and increases the idle time of many resources. Yet, as soon as the projects start, early or late task completions create new conflicts between the projects. It seems you cannot win, even with longer and longer project plans. The CCPM multi-project solution (1) schedules individual projects by CCPM without conflict, (2) sequences the projects according to a fixed point (or strategic resource or task), (3) manages all projects with BM, and (4) assigns resources to projects (and tasks) as needed to benefit the overall organization.

³[Figure 33-4](#) follows an approach first used by Alan Barnard in his presentation, “Insights and updates on the theory of constraints thinking processes” at the first Theory of Constraints International Certification Organization (TOCICO) Conference held in Cambridge, England in 2003.

⁴See, for example, Section II on Critical Chain and Section III on Drum-Buffer-Rope and the Distribution/Replenishment solution.

⁵See [Chapter 14](#)—Resolving Measurement/Performance Dilemmas.

[6](#) © E. M. Goldratt used by permission, all rights reserved.

[Z](#)Note: The use of CCPM here shows how the overall project for ideas should be planned. That is, there is a 50 percent project buffer protecting the critical chain and 50 percent feeder buffers for the feeding chains. However, the concern of the complex organization is to meet the expected delivery dates (promises) from one group or department to another. The project “ideas” can be managed as a whole according to CCPM methods, but the complex organization must deliver according to plan or there will be major disruptions to the whole system. This will become more obvious later in this chapter. The solution for the complex organization is an overriding measurement system that will govern the flow of all groups and departments, not just the project, production, and distribution worlds that are part of the complex organization.

⁸TDD is discussed in *The Haystack Syndrome*, [Chapter 24](#) (Goldratt, 1990); *The Theory of Constraints Journal*, 3:17-18 (Goldratt, 1988a); TOC Insights for Distribution, Parts 10 and 11, Measurements of Execution (Goldratt et al., 2006); and in the *TOCICO Dictionary* (Sullivan et al., 2007). (©TOCICO 2007, used by permission, all rights reserved.)

⁹Here we see that the groups, departments, and organizations receive TDD assessments based upon their delivery reliability. We also see that the product or project can report TDD as well. That is, the TDD assessed to a unit is also recorded with the product or project. This is not a double accounting, but a record of what group or department had unreliable delivery and also what product caused it. Having the TDD measured in this way helps management assess which organizations need help and which products/projects need improvement as well as which groups need help.

10 While recovering buffer is important in process flows, it is neither rewarded nor punished. Rewarding or punishing buffer penetration or buffer recovery creates the wrong behavior; task estimates become inflated and aggressive plans are lost.

[11](#)Shippers Supply Company has used TDD for over three years. They only had to add one data element to their existing database in order to calculate the TDD. The daily TDD report is enormously effective in reducing what was an unmanageable stock level. They use the TDD report to expedite late work. They use a Pareto analysis to focus their improvement efforts. Customer service is very happy; now, customers rarely complain. The TDD numbers alert management in advance so they can take actions to fix problems before they affect delivery (Johnson, 2009).

¹²[Figure 33-16](#) was first drawn by John Thompson (2009).

[13www.tocico.org](http://www.tocico.org)

[1](#)The S&T tree is one of a number of tools of the TOC Thinking Processes, which can be read about in Chapters 24 and 25 of this book. The S&T tree includes both sufficiency-based logic and necessity-based logic, which are described in those chapters and in more detail later in this chapter. The S&T tree should be written after the CRT and FRT have already been developed.

²In TOC costs are classified as totally variable costs, Operating Expense, and Investment.

[3](#)The full explanation and logic underlying finding this win-win solution is provided in Goldratt (2008a).

⁴The trees are available in the member section of the TOC International Certification Organization (TOCICO) Website at www.tocico.org and as part of a useful software program named Harmony for creating S&T trees at www.goldrattresearchlabs.com. Note that the most up-to-date versions of the S&T trees are automatically included in Harmony. The full Retailer S&T tree is not presented here due to space limitations. The missing steps of the full S&T tree can be downloaded at the above sites and are read in a similar manner as presented in our discussion here.

[5](#)This win-win solution is explained in *The Choice* (Goldratt, 2008a) in Chapters 2, 8, and 10.

6 There are two buy-in processes in TOC, which are referred to as the plus and minus-minus processes.

[Z](#)These solution elements were presented in The Goldratt Webcast Program on Project Management (Goldratt, 2008b).

[8](#)Note that the combination of the Retailer and CG S&T trees provides the win-win solution for both retailers and suppliers that is explained in *The Choice* (Goldratt, 2008a).

⁹It is recommended that anyone who is interested in this S&T tree review the program that Dr. Goldratt facilitated, which provides a full explanation of this S&T tree. It is available on DVD, titled “The Goldratt Webcast Program on Project Management,” at www.toc-goldratt.com.

[10](#) My suggestion for learning more about the literature on strategy is to review Thompson, Strickand III, and Gamble (2008).

[11](#) Suggestions for how to implement this approach are described in the TOC Insights into Distribution and Supply Chain, which is available at www.toc-goldratt.com.

¹²See the TOC Insights into Operations, which can be purchased at www.toc-goldratt.com, for a good explanation about why local efficiencies are not a good measure of performance.

[13](#) He pointed out that two additional usages of S&T trees are for project management (for choosing the project and determining its content) and as an organizer of knowledge.

¹A discussion of complexity is presented in Goldratt, E. M. 1987. *Theory of Constraints Journal* 1(5) [Chapter 5](#)—“How complex are our systems?” New Haven, CT: Avraham Y. Goldratt Institute. (© E. M. Goldratt used by permission, all rights reserved.)

²A form of this cloud was presented in Goldratt, E. M. 1999. Goldratt Satellite Program. Session 2 Finance and Measures. (© E. M. Goldratt used by permission, all rights reserved.)

³Goldratt, E. M. 1990 . *What's This Thing Called Theory of Constraints?* Croton-on-Hudson, NY: North River Press.

⁴In the *APICS Dictionary* (Blackstone, 2008, 2), *common cause* is defined as “Causes of variation that are inherent in a process over time. They affect every outcome of the process and everyone working in the process. Syn: random cause. See: assignable cause, assignable variation, common cause variability.” (© APICS 2008, used by permission, all rights reserved.)

Special or assignable cause, in the *APICS Dictionary* (Blackstone, 2008, 7) is defined as “A source of variation in a process that can be isolated, especially when its significantly larger magnitude or different origin readily distinguishes it from random causes of variation.” (© APICS 2008, used by permission, all rights reserved.)

[5](#)Project management solutions are discussed in Section II and logistics solutions are discussed in Section III of this Handbook.

⁶The *TOCICO Dictionary* (Sullivan et al., 2007, 15) defines *critical chain* as “The longest sequence of dependent events through a project network considering both task and resource dependencies in completing the project. The critical chain is the constraint of a project.

Usage: The critical chain plus the project buffer determines the lead time for the project. If no resource contention exists, then the critical chain would be the identical to the critical path.” (© TOCICO 2007, used by permission, all rights reserved.)

[1](#)The *TOCICO Dictionary* (Sullivan et al., 2007, 18) defines *Drum-Buffer-Rope (DBR)* as “The TOC method for scheduling and managing operations. Usage: DBR uses the following: 1. the drum, generally the constraint or CCR, which processes work in a specific sequence based on the customer requested due date and the finite capacity of the resource; 2. time buffers which protect the shipping schedule from variability; and, 3. a rope mechanism to choke the release of raw materials to match consumption at the constraint.” (© TOCICO 2007, used by permission, all rights reserved.)

⁸The *TOCICO Dictionary* (Sullivan et al., 2007, 17) define *distribution/replenishment solution* as “A pull distribution method that involves setting stock buffer sizes and then monitoring and replenishing inventory within a supply chain based on the actual consumption of the end user, rather than a forecast. Each link in the supply chain holds the maximum expected demand within the average replenishment time, factored by the level of unreliability in replenishment time. Each link generally receives what was shipped or sold, though this amount is adjusted up or down when buffer management detects changes in the demand pattern.

Usage: The largest amounts of inventory are held at a central warehouse where the variation in demand is the least. Smaller amounts of inventory are held and are replenished frequently at the end consumer location where variation in demand is the greatest. Throughput dollar days and inventory dollar days are measures used to judge the reliability and effectiveness, respectively, of each link in the chain. Transfer pricing is not used.” (© TOCICO 2007, used by permission, all rights reserved.)

¹The *APICS Dictionary* (Blackstone, 2008, 136) defines *Takt time* as “Sets the pace of production to match the rate of customer demand and becomes the heartbeat of any lean production system. It is computed as the available production time divided by the rate of customer demand. For example, assume demand is 10,000 units per month, or 500 units per day, and planned available capacity is 420 minutes per day. The *Takt time* = 420 minutes per day/500 units per day = 0.84 minutes per unit. This *Takt time* means that a unit should be planned to exit the production system on average every 0.84 minutes.” (© APICS 2008, used by permission, all rights reserved.)

²Used by permission of Eliyahu M. Goldratt © Eliyahu M. Goldratt.

³The TOCICO *Dictionary* (Sullivan et al., 2007, 41) defines *relay runner* as “*The process of applying a focused effort to complete a task and hand it off immediately to a resource waiting and prepared to take the hand-off in critical chain project management.*”

Usage: Some people use relay runner interchangeably with road runner in an operations environment. (© TOCICO 2007, used by permission all rights reserved.)

⁴The APICS Dictionary (Blackstone, 2008, 93) defines order point system as “The inventory method that places an order for a lot whenever the quantity on hand is reduced to a predetermined level known as the order point.” Two order point systems are used: the min-max system and economic order quantity system (EOQ). The min-max system (83) is “A type of order point replenishment system where the “min” (minimum) is the order point, and the “max” (maximum) is the “order up to” inventory level. The order quantity is variable and is the result of the max minus available and on-order inventory. An order is recommended when the sum of the available and on-order inventory is at or below the min.” The EOQ (43) system is defined as “A type of fixed order quantity model that determines the amount of an item to be purchased or manufactured at one time.” (© APICS 2008, used by permission, all rights reserved.)

¹This particular case study was the topic of an Industry Week Webinar, which they later claimed was their highest attended ever.

[2](#) Recall that Jonah is a character in *The Goal* (written by Eli Goldratt) that Socratically leads a person to discover the answer to a problem.

¹Some of the materials in this chapter are drawn from Cox, Blackstone, and Schleier (2003, [Chapter 17](#)).

²One of the authors started teaching the TP to his college students based on personal productivity and time management dilemmas. His thinking was that it was easier to teach a new methodology if the student was familiar with the subject matter on which the methodology would be used. This approach proved to be of significant value to a large number of students. The story of one such student is provided later in this chapter.

³The *TOCICO Dictionary* (Sullivan et al., 2007, 11) defines a *chronic conflict* as: “A contentious situation that has continued to exist for a prolonged period of time. Opposing sides have been justifying their perspective through selective requirements and prerequisites for so long that both sides become entrenched in their own beliefs to the point that neither side can see how to break the conflict without suffering a significant loss.

Usage: Breaking a chronic conflict requires understanding the opposing perspectives. This understanding can lead to the surfacing of hidden assumptions underlying entity relationships that are often the key to creating a breakthrough solution. The solution to a chronic conflict requires one side to offer up a problematic (from their perspective) injection and the other side to somehow eliminate any of the undesirable aspects of the proposed injection using negative branch reservations (NBRs).” (© TOCICO 2007, used by permission, all rights reserved.)

⁴After a few iterations of this cloud, my son realized that I was concerned with his security. Therefore, he would seek out answers to the anticipated questions before we started our dialog. Occasionally, I would have a question where his response was, “Good question! Let me find out what the answer is and get back to you.”

⁵This problem was studied by a number of different MBAs over the past decade; for a detailed analysis of this problem and to see a different way of building a future reality tree see: Cox, Mabin, and Davies (2005).

⁶In [Chapter 16](#), Barnard and Immelman use Ackoff's terminology of Errors of Commission (doing what you shouldn't do) and Errors of Omission (not doing what you should do) as causes of grave problems. These same categories fit well what we are describing here.

[Z](http://www.sheilat.com/keynote.htm)Sheila not only made the 2000 and 2004 Olympic teams in the triathlon, but also she switched sports and competed in the pentathlon in Beijing in the 2008 Olympics. She is the only woman ever to compete in four Olympics in three different sports. To read more of her amazing story, go to: <http://www.sheilat.com/keynote.htm>