Biology & Philosophy (2019) 34:2
https://doi.org/10.1007/s10539-018-9650-2

BOOK REVIEW ESSAY

Review of Other Minds: the octopus, the sea and the deep origins of consciousness

Peter Godfrey-Smith, Farrar, Straus and Giroux, NY, 2016

Daniel Dennett[1]

Received: 2 January 2018 / Accepted: 20 October 2018

Some years ago there was a workshop on convergent evolution, starring two species: the octopus and us human beings. Mollusk and Man, as one used to say. What next, sunflowers and sidewinders? (They both keep track of the sun, after all.) But this is no joke; there are striking parallels between the minds of human beings and the minds of octopuses, and studying them carefully and open-mindedly is a great way of understanding what minds are or might be. Our most recent common ancestor with the octopus was a tiny worm that lived hundreds of millions of years ago. As Peter Godfrey-Smith says,

> Cephalopods are an island of mental complexity in the sea of invertebrate animals. Because our most recent common ancestor was so simple and lies so far back, cephalopods are an *independent experiment* in the evolution of large brains and complex behavior. If we can make *contact* with cephalopods as sentient beings, it is not because of a shared history, not because of kinship, but because evolution built minds twice over. This is probably the closest we will come to meeting an intelligent alien. (p. 9)[1]

This passage packs a philosophical wallop. Philosophers are fond of imagining "Martians" and "zombies" and other alien intelligences, but their imaginations are typically not up to the task. Whatever they hold constant (without realizing it) is

[1] There is an even more alien form of life that has a nervous system, nowhere near as intelligent as an octopus, that is now recognized as profoundly different (at the level of neuromodulators and proteins and genes) from all other animals: the lowly ctenophores, a distant relative of jellyfish with which it is often confused. See https://aeon.co/essays/what-the-ctenophore-says-about-the-evolution-of-intelligence?utm_source=Aeon+Newsletter&utm_campaign=0b9c554b15-EMAIL_CAMPAIGN_2017_11_27&utm_medium=email&utm_term=0_411a82e59d-0b9c554b15-69421497.

✉ Daniel Dennett
Daniel.Dennett@tufts.edu

[1] Tufts University, Medford, USA

Published online: 01 December 2018

likely to be ill-motivated, and the variations that do occur to them (x-ray vision, astronomically capacious factual memories, direct experience of cosmic rays,…) seldom manage to get at the truly interesting and important differences that we can discover if we actually do the science. Moreover, doing the science can enlarge and enhance our "direct" experience of other minds—if we make *contact* with them, as he says. This is not a trivial supplement of empirical fact. Our sense of "what it's like" to be *us* is profoundly shaped by the contact we make with each other, and—for better or for worse—our convictions about *whether* "it is like something" to be an octopus, are apt to be grounded by, perhaps overwhelmed by, actual interactions with octopuses. If it is true—as it seems to be, on Godfrey-Smith's eloquent and detail-rich account—that you cannot help but arrive at the confident opinion that octopuses are conscious once you've engaged with them, this tells us at least this: before you even begin to study the interior complexities of these amazing creatures, your default assumption will be: they are, like us, conscious in some special way, not merely sensitive (like bacteria, like plants,…) and probably not merely sentient (like our shared ancestor worm, like ants,…) but—in some way still to be characterized—conscious. *Subjective experience* is Godfrey-Smith's preferred term for this elusive quarry, and he claims that the cephalopods are good model organisms to study, as biologist say, since they share *so little* with us: they are not apes, not primates, not mammals, not warm-blooded, not even vertebrates.

Another punch in the passage quoted above reminds us that a good way to study minds sets aside, to the extent possible, the Wittgensteinian "way of life" we share in our linguistic communities, the commonalities that are apt to confound our thinking with parochiality. We do *not* share a history, or a close kinship, with the octopus and yet it has a mind, as we can see if we just look. These philosophical messages interact: they point to the almost irresistible intuition that can be pumped by what might be called *mere behavioral tempo and rhythm*: if cephalopods moved in the clunky way of most existing robots, then in spite of the manifest purposiveness of their motions, it would be quite comfortable to suppose that they were some kind of zombies, marine robots with eight or ten appendages. (Cog, the humanoid robot developed by Rodney Brooks and his team at MIT some years ago, never approached consciousness, but it moved its arms and eyes and head with such humanoid vivacity and even grace that naïve observers often blurted out loud their startled conviction that it was conscious.)

Godfrey-Smith is well aware of the potency of such intuitions and the pitfalls confronting the imaginations of researchers, and he scrupulously describes and explains the discoveries he has made in his years of underwater research as an amateur (in the best sense) marine biologist. His conclusions are informed, first, by a careful review of the evolutionary origins of the cephalopods, and then by a survey of the amazing cognitive talents they have developed. As he reads the evolutionary record, "There's a smooth transition from minimal kinds of sensitivity to the world to more elaborate kinds, and no reason to think in terms of sharp divides" (p. 77). Similar selection pressures in widely separated lineages have driven similar solutions to fundamental problems of staying alive, not always in the same order in each lineage, so we have a patchwork of different talents exhibited in different phyla. Cephalopods are mollusks, like clams and oysters, and part

of the evolutionary message is clear: if you are going to be a mollusk without a shell, you had better be smart—but even this policy has exceptions, as Godfrey-Smith reminded me: slugs are not celebrated for their cleverness, while the lowly *Aplysia* is prized by scientists for the simplicity of its nervous system.

Octopuses turn out to be quick studies: they are excellent learners, but have shockingly short lifespans—one or two years, in general, with four years as the outer limit. Why invest so heavily in learning if you're not going to be able to use your acquired skills for long? Why have a short life span in the first place? These and other 'why" questions have answers—to some extent worked out and confirmed—in adaptationist terms: there are *good reasons* to be designed this way or that way. In a lucid summary and extension of evolutionary reasoning by Peter Medawar, George Williams and William Hamilton, Godfrey-Smith shows how (1) late-acting deleterious mutations accumulate (as byproducts, not with their own purposes) to create senescence, and how (2) the *rational policy* of animals where the risks of predation in adulthood are high is to "go for broke" (p. 171) and make a large early investment in reproduction instead of distributing the reproductive project over many years or seasons.

> The lifespan of different animals are set by their risks of death from external causes, by how quickly they can reach reproductive age, and other features of their lifestyle and environment…. Early cephalopods had protective external shells… Then the shells were abandoned… it gave cephalopods their outlandish, unbounded possibilities. (pp. 72–73)… They have the large nervous systems because of what those unbounded bodies make possible and the need to hunt while being hunted; their lives are short because their vulnerability tunes their lifespans. (p. 174)

This brief summary cannot do justice to the patience and ingenuity with which Godfrey-Smith considers and dismisses alternative reasons, and the way he accumulates the empirical facts supporting his argument. The whole explanatory enterprise is biology as *reverse engineering*, and if you want to *understand* consciousness (as contrasted with just staring gob-smacked at it and declaring it a mystery), you must ask and answer these fundamental questions, because consciousness is clearly expensive, both in terms of the R&D that has gone into its emergence over millions of years of evolution, and metabolically. Big complex brains are energetically expensive, even though they have been optimized for energy efficiency for eons. That is part of what makes the decentralized, distributed computational architecture of the cephalopods so fascinating: it turns out that there is more than one way of designing a nervous system that can think ahead effectively.

Thinking ahead is the main task of nervous systems: preparing the organism to make timely responses to its multifarious needs and opportunities. Life is an overlapping series of arms *races,* and just as computers have relentlessly increased their fundamental speed of processing, nervous systems have found faster and cleverer ways of looking ahead, taking informed chances, cutting corners, pruning out bottlenecks and other slowdowns, in addition to simply ramping up the

speed of signaling within the systems. In order to do what, exactly? What product or outcome is to be achieved by all this engineering? Here is where many theorists of consciousness make a simple mistake: they stop one (giant) step too early. They devote great labor investigating the upward or inbound paths "to consciousness" and when they think they have reached their target, they stop and declare victory. They don't ask the Hard Question: 'And then what happens?" (Dennett 1991, p. 255) But it is answers to the Hard Question that are the only hope of demystifying consciousness. What does consciousness permit an organism to do, and how?

One still powerful philosophical tradition would say that this question is already a mistake, a "behaviorist" mistake of some sort:

> Many years ago Thomas Nagel used the phrase *what it's like* in an attempt to point us toward the mystery posed by subjective experience…. The term "like" is misleading here, as it suggests that the problem hinges on issues of comparison and similarity—*this* feeling is like *that* feeling. Similarity is not the issue. Rather, there *is a feel* to much of what goes on in human life…. That's what has to be understood. But when we take an evolutionary and gradualist perspective, this takes us to strange places. How can the fact of life feeling like something slowly creep into being? How can an animal be halfway to having it feel like something to be that animal? (p. 78)

Is this just a rhetorical question? For some philosophers and psychologists, it is, but Godfrey-Smith doesn't fall into that trap. He sets out to answer the question. There being *a feel* is commonly seen as sharply contrasting with being able to do something (behave somehow) in reaction to that feel; turning our attention to the reactions seems to be turning *away* from the *subjective experience itself*. But this is just wrong. What *gives* the subjective experience itself the *content* it has, what makes it the subjective experience it is, must be what it *does to* the subject—typically what it does *for* the subject, but also, what disruptions if any it causes in the organism's normal systems of self-maintenance. (Not just immediate effects either; the organism may salt away some effect of its current subjective experience for use in modulating some much later activity).

> What you'll do next is affected by what you're now sensing; and also, what you'll *sense* next is affected by what you now *do*…. We know that explicitly and can talk about it, but their intertwining also affects in more fundamental ways how things feel, in quite a raw sense of 'feel.' (p. 80)

Godfrey-Smith doesn't use Gibson's increasingly popular term, *affordance*, but he has a lot to say about how these tightly and automatically intertwined transitions between perception, motivation, expectation and preparation for action provide the contents, the flavors, the gists of subjective experience. (This is a close cousin of my way of putting the point (e.g., Dennett 2017) that the brain has developed a *user-illusion* that simplifies its task of generating anticipations and timely guidance of actions. This feeds quite naturally and directly into the various global workspace and working memory theories of consciousness, and Godfrey-Smith

appreciates the strengths of these haltingly converging approaches, but he sees them, which he usefully dubs *latecomer* theories, as too anthropocentric (or just mammalcentric): they fail to find room for something beyond mere appropriate sensitivity but short of self-reflection in, say, fish and invertebrates. "Do you think that those things (pain, shortness of breath, etc.) *only feel like something* because of sophisticated processing in mammals that has arisen late in evolution? I doubt it." (p. 93) He has some tentative ideas about what he calls a *transformation* view of subjective experience: "In this picture, there are early and simple forms of subjective experience that are then transformed as evolution makes nervous systems more complicated." (p 95) As complications are added (gradually, of course, in rudimentary and then less rudimentary forms), "new capacities—such as sophisticated kinds of memory—are added which have a subjective side." (p. 95). Can he really be suggesting that a *pain that can be remembered* (for instance) is more like a *feeling* than a pain (would it *be* a pain?) that had no *sequelae* at all? Yes, and it is just such useful increments that must be posited to get us out of the self-defeating pit of what might be called *feeling-trance*: the paralysis that prevents one from asking oneself "and then what happens?" And once again, there are intermediate, smaller steps to be considered: it is one thing to be able to have, and reflect on, an *episodic memory* of a past pain and some of its circumstances, and another thing to have one's dispositions and preferences regarding those circumstances adjusted aptly in ways likely to avoid "pain" (if it is pain) in the future. Yes, an animal can be "half way towards having it feel like something." Later, with the elaboration of "a more definite sense of self" we reach "something closer to consciousness. I don't see that as a single definite step." (p. 97)

Many philosophers say that they cannot imagine any intermediates between the unconsciousness of, say, a brick, and their own sort of consciousness, and then they draw a "conclusion" from this noted inability: the light of consciousness is either on or off. Godfrey-Smith shows them, by example, that if they would just try harder, by actually considering some plausible details, they might succeed.

In chapter 6, "Our minds and others," his evolutionary vantage point lets Godfrey-Smith draw attention in a quite unified way to the various competing proposals amongst theories of consciousness recently put forward by philosophers, neuroscientists, computer scientists and others. The adherents of these positions tend to see themselves as at war with all the rivals, but Godfrey-Smith's framework allows him to find valuable contributions to his overall scheme without appearing to be merely a cherry-picker, an opportunistic Polyanna who can find a little bit of good in everyone. The various *global workspace* ideas (of Baars, Dehaene, Carruthers, Prinz, and me) fit in handsomely as long as we treat them as describing a *broadcast* system between otherwise non-communicating parts; the *higher-order thought* theory (of Rosenthal and others) is acceded a valuable role in the *latecomer* varieties of consciousness (I think it has other virtues, but that is a topic for another occasion); the *integrated information* idea (of Tononi, and now Koch) gets a constructive gloss, as does Hofstadter's idea of *strange loops* (though without attribution to Hofstadter). He considers imaginatively the particularly explosive role of adding language to the mix of capabilities, but carefully points out that at least crude versions of such expressive expansion may exist in other species, a possibility now being explored experimentally. And he adds some useful discussion of *efference copy* as a source for silent inner speech (though in an endnote he *contrasts* this

with my account, with which it is entirely consonant). In an academic monograph, this 20-page roundup and evaluation of rival ideas would take hundreds of pages and even more references, but this book is intended for a wider readership, and its compactness serves to introduce the ideas vividly, which is the main point. (His decision to pack all the bibliographic citations—which are good—into endnotes unmarked in the text may give the mistaken impression that he is unaware of many of the sources for his ideas. This favoring of graceful, easy reading over Giving Proper Credit may annoy some professionals, but their students will pick up the ideas with less toil, and any curiosities engendered are likely to be well rewarded with further reading found in the endnotes.)

One of the too-brief discussions in the book is his acknowledgment of Simona Ginsburg and Eva Jablonka's ideas (2007) about the gradual evolution of experience, but this gap can be filled to a considerable extent by Jablonka's presentation (following Godfrey-Smith's) and my commentary on both of them at the NYU conference on animal consciousness, November 17–18, 2017, on line at https://wp.nyu.edu/consciousness/animal-consciousness/.

There is much else that is fascinating and important in this small volume, not least the amazing details of cephalopod life, and the surprising recent discoveries about convergent or parallel evolution in multiple branches of the Tree of Life. The details do matter. How can the octopus control its astonishingly swift and accurate changes of color when it is, apparently, colorblind? It turns out that its skin is itself light-sensitive. (What is it like to see with your skin? Wrong question, since it is only *sorta* seeing, and there need be no single central Subject that enjoys the show.) How do cuttlefish use their ability to create moving patterns of color on the "living video screen" (p. 160) on their sides? If they are signaling, who is interpreting the messages, and what is their function? These are still largely unanswered questions, but Godfrey-Smith whets the reader's appetite for them, and points to the future work that will answer them.

Socrates famously described himself as a midwife helping others to give birth to wisdom, and Aristotle's compendious knowledge of nature and boundless curiosity permitted him to assist at the birth of many important scientific ideas (a few of them stillborn). The great philosophers of the seventeenth and eighteenth century engaged energetically with the science of their day and contributed to the shape science assumed thereafter—often but not always positively. Today, we are witnessing something of a payback, with science providing—for philosophers who will attend—much needed conceptual breakthroughs, imagination-prostheses, and even methods of thinking that are dismantling ancient philosophical logjams, turning traditional "mysteries" into soluble problems. Peter Godfrey-Smith provides an excellent example of what can be accomplished by a scientifically educated philosophical imagination.

References

Dennett D (1991) Consciousness explained. Little, Brown, New York

Dennett D (2017) From bacteria to bach and back: the evolution of minds. W.W. Norton & Co., New York

Ginsburg S, Jablonka E (2007) The transition to experiencing: I. Limited learning and limited experiencing. Biol Theory 2:218–230