Dennett, Daniel C. 2017. *From Bacteria to Bach and Back: The Evolution of Minds*.
New York: W. W. Norton. xviii, 496 pages.
Hardback $28.95; Kindle Edition $14.14.

Ronald de Sousa

Over close to half a century, Daniel Dennett has been the model of a scientifically informed philosopher exploring the mind as a product of evolution. His latest book is a magisterial summa*: he is "going for the whole story this time" (16).

The key to Dennett's reconstruction of our species' progress to its unique form of intelligence is the concept of *competence without comprehension.* Every living thing—from bacteria to Bach—persists and reproduces by dint of clever tricks. At first, all of these tricks are performed without *comprehension*. That is Darwin's "inversion" of the commonsense assumption that competence derives from comprehension. Only after millions of years of such mindless competence has this process culminated in the appearance of creators like Bach, whose top-down designs are based on a rich understanding of what he wanted and how to achieve it. And thanks to a second "inversion," this time from Alan Turing, human engineers have constructed machines whose increasingly stunning competences, in their turn, require no comprehension.

Dennett boldly insists that the products of natural selection literally embody design, just not the kind that required a designer. Many Darwinians deplore such unqualified use of teleological language in reference to biological functions, fearing that it might give comfort to the "intelligent design" crowd. Actually, however, it pulls the rug from under them. Indeed! Dennett can calmly assert, the biosphere is rich with marvels of adaptation! They are due to the demonstrable reality of competences that result,

bottom up, from natural selection, without benefit of intelligence.

That way of talking plays a crucial role in Dennett's overarching project. It prepares the reader for the insight that even in our most intelligent performances, we make use of competences that we don't understand and have no need to understand. Even a Bach, in his most explicitly intelligent contrivances, will have relied not only on others' crafts, but also on innumerable subpersonal routines. Such routines enable each of us, for example, effortlessly to perform each explicit step in a chain of conscious reasoning, such as writing down the result of an arithmetical operation. When we realize that our free and creative use of language involves competences we had no part in fashioning, we should the more readily concede that the difference between fully intentional conscious design and uncomprehending competences is a matter of degree.

The key role of competence without comprehension extends to the evolution of culture. To illustrate this, Dennett makes use of Peter Godfrey-Smith's neat concept of "Darwinian spaces," a convenient way to represent the bearing of parameters that complement and refine Darwinism's standard triad of variation, heredity, and differential reproduction. In one such space, the 0,0,0 corner is inhabited by purely Darwinian processes, involving zero comprehension, bottom-up design, and random rather than directed search. At the opposite corner lies the ideal conscious design, scoring a maximal 1 in comprehension, top-down

conception, and consciously directed search. In practice, all "real cultural phenomena occupy the middle ground, involving imperfect comprehension, imperfect search, and much middling collaboration" (144). Figure 7.5, on which the sentence just quoted is a comment, doesn't actually have a dimension of collaboration. Implicitly, then, the comment equates a high reliance on collaboration with "bottom-up" construction. Why that might seem reasonable is not explored until the last chapter; so perhaps we can take this as an illustration of Dennett's own concept of a *thinko*—the semantic analogue of a typo (225). The thinko doesn't affect the main point about culture, which is that it will never, any more than individual consciousness, attain a perfect score on any of those dimensions:

> Human culture started out profoundly Darwinian, with uncomprehending competences generating various valuable structures in roughly the way termites build their castles. Over the next few hundred thousand years, cultural exploration of Design Space gradually de-Darwinized … becoming a process composed of ever more comprehension. (282–83)

The main instrument of that process of "de-Darwinization," lifting us from the Darwinian base towards full top-down comprehension-based design, was the spread of *memes*. "Memes" refers to cultural "ways" that "go viral" among human minds, bringing Darwinian mechanisms into the heart of culture. That has made them seem threatening to the vaunted autonomous creative capacities of intelligent humans. The threat is real, in part: the spread of memes, like that of genes, isn't necessarily good for their hosts. In the service of their own propagation, genes can be deleterious to their carriers. Similarly, a meme's aptitude to grip onto human minds is only weakly related to epistemic, moral, or aesthetic standards we profess.

As an objection to memes, the demotion of human creativity they appear to entail is feeble. It amounts to little more than the expostulation that "it would be too sad if it were true." But

to many critics, a more significant defect is the term's vagueness. A meme can be almost anything: a phoneme, a word, a dance, a cultural practice. The extent to which memes are replicable, some protest, varies far too much to sustain the analogy with genes.

For Dennett, however, that is not a problem. In digital systems of representation, every sign is identified as one element or another from a finite set. None is ever anything in between. So fidelity in replication is indeed a crucial feature of such systems. But digitality itself admits of degrees. As Dennett notes, "hearsay—stenography—vinyl records—DNA—digital file" form a continuum of increasingly accurate replicability (135). Degrees of fidelity affect the durability of cultural practices, but lower degrees can also be useful. Words are at the very high end. They are truly digital. Most importantly, words form, once we start to use and reflect on them critically, the essential "cranes" that lift us, in a cascade of metalevel questions, ever higher into the dimension of comprehension.

Although Dennett's championing of the concept of meme has met with a good deal of resistance, what many have found most baffling is his view of consciousness as "user illusion." If consciousness is an illusion, who is being deceived? This is where the pull of "Cartesian gravity" is strongest: surely I can be deceived only if I am the conscious *subject* of my illusion!

On this, Dennett has plausibly been suspected of mischievous teasing. Actually, however, he nowhere denies the existence of experience. His point is rather that, like the computer display that grounds the "user illusion" metaphor, my experience provides almost no clue to the mechanisms that give rise to it. As the elaborate description of the role of competences without comprehension has shown, our own thinking in some ways resembles GOFAI (Good Old-Fashioned Artificial Intelligence) programs written in a high-level language. It makes constant use of routines and subroutines of which we know and care about only

the outcome—the content of the user illusion: "Our access to our own thinking, and especially to the causation and dynamics of its subpersonal parts, is really no better than our access to our digestive processes" (346). If comprehension, directed search, and top-down design admit of degrees, then so does the consciousness that is constituted by those mental activities.

Thus primed, we are less likely to believe that consciousness is an all-or-nothing affair, with which some creatures are endowed, but others not. In the evolution of our phylum, or the path from a zygote to an adult Bach, when exactly did consciousness pop up?

Dennett's rejection of qualia, the *what-it's-like* character of experience, has been equally controversial. As conceived by most of those who use the term, a red-and-stripey quale is required to explain my conviction that I see a red stripe. Maybe it is felt to be required because neither the brain process nor my conviction is either red or stripey. But if the brain processes cause—or constitute—the conviction, "the postulation of qualia is just doubling up the cognitive work to be done" (363).

Still, some might balk, removing qualia as useless intermediates between a perceptual state and its causes doesn't overcome Cartesian gravity. For am I not, ineliminably, the *subject* of my conviction that I perceive a red stripe? Even if the quale isn't required to explain anything, is it not still perhaps an epiphenomenal, or causally inert, object of my experience?

If qualia were causally inert, however, they would be invisible to natural selection. Unlike consciousness as the elaborate compound of competences at various stages of comprehension, the advantages of which can obviously be targets of natural selection, qualia would be a gratuitous add-on. They might have spread, but never to fixation. Some of us, then, might be "zombies"—beings that behave in all respects exactly like a normal person, but for whom there is no *what-it's-like* to anything. This is where Dennett performs what we might call a zombie jiujitsu move. If you can make sense of the concept of zombies, you can imagine a being convinced that it is seeing blue yet lacking the quale that is either the cause or the epiphenomenal object of that conviction. But if this possibility is admitted, then how can you be sure you are not a zombie?

> "But I know that I am not a zombie!" No, you don't. The only support for that conviction is the vehemence of the conviction itself, and as soon as you allow the theoretical possibility that there could be zombies, you have to give up your papal authority about your own non-zombiehood. (357)

If qualia are epiphenomenal, then zombies are possible. And if zombies are possible, you might just be one.

In the last chapter, Dennett considers the astonishing rise of machines capable of indefinite progress on the basis of unsupervised learning. Chess, *Jeopardy*, and Go have all confounded the skeptics who thought them forever out of reach for machines. Go's possible moves are in the region of 10^{170} (for comparison, all the atoms in the universe are thought to number around 10^{80}). So it cannot be solved in practice by exhaustive search in real time. Given these triumphs, readers may be surprised to find that Dennett does not think machines are on the verge of achieving true comprehension. The reason is that although some self-monitoring can be programmed into computers, only genuinely self-critical organisms, and perhaps only those equipped with bodies and goals of their own, can achieve true comprehension: "The kind of comprehension AI systems are currently exhibiting—and it is becoming breathtakingly competitive with the best human comprehension—is also parasitic, strictly dependent on the huge legacy of human comprehension that it can tap" (397). That means that as far as we can tell (but we are only human), "in the next fifty years" (399) AI is unlikely to become an autonomous deliberative agent, whose competences attain genuine comprehension. In that regard, however impressive

their performance, they remain *back* on a par with bacteria.

Dennett hints at the ironic possibility that comprehension may have peaked. Big science now demands massive collaboration. The paper reporting the Higgs boson detection was signed by more than three thousand authors. And unlike the artisans of GOFAI, who knew what their machines were doing because they programmed them, the designers of AlphaGo Zero have only the most general understanding of what their creation is creating. A wistful corollary is that just as women are poised to explode the myth of exclusively male genius, they may miss out on that status because no man or woman will ever again be Newton, or Einstein.

This last point raises questions of gender, first broached early on in a curiously explicit provocation (23–24). A reader who took the bait may not be appeased when the issue comes up again. Dennett dismisses those who "resist the quite obvious free-floating rationale for why it is females that do the evaluating and males that do the strutting … asymmetrical parental investment" (134n). That logic is neat, but it is swamped by the diversity of sex-gender arrangements found in nature. As Joan Roughgarden has written, "The biggest error of biology today is uncritically assuming that the gamete size binary implies a corresponding binary in body type, behavior, and life history. … Gender, unlike gamete size, is not limited to two" (2004, 26–27). A good reason to "resist" what Dennett endorses as "quite obvious" is that much science of sex has proceeded on the basis of groundless "commonsense" assumptions about gender (Jordan-Young 2010). These then reemerge as "conclusions" about "natural" gender differences that are used to support the political backlash against women's equality.

Another quibble concerns what Dennett describes as his "conversion" from his long-held "homuncular functionalism"—the view that the brain is organized in a hierarchy of increasingly simple functional units. The nature of this change of mind, for which he credits Terrence Deacon (2011), is frustratingly obscure. He adduces neuronal plasticity: "Neurons … play more enterprising and idiosyncratic roles than the obedient clerks I was imagining them to be" (162). Sometimes they go "feral" (171). But that hardly warrants a "conversion," and Dennett still thinks his idea was "on the right track." What else is there? Perhaps the appendix provides a clue, where he alludes to Adrian Thompson's "evolving hardware" (423). When a machine was set up to physically rewire itself, it exploited physical glitches, deviating from strict digitality. As a result, the circuits "worked using principles that human designers would never dream of employing" (Bentley 2001). Tantalizingly, Dennett says he has "often spoken about" this work but has "not discussed it in print" (423). It would have been good to clarify its significance in the context of Dennett's "conversion."

Quibbles aside, this is a fine summa, not least for its captivating side trips. "Semantic" versus "Shannon" information; the apparent Lamarckism of Baldwin selection; the origins of language; the relation between the manifest and scientific images; the promise of CRISPR technology: all these and many more topics are touched on, with Dennett's usual wit and polymathic panache.

WORKS CITED

Bentley, Peter J. 2001. *Digital Biology: How Nature Is Transforming Our Technology and Our Lives*. New York: Simon & Schuster.

Deacon, Terrence W. 2011. *Incomplete Nature: How Mind Emerged from Matter*. New York: W. W. Norton.

Jordan-Young, Rebecca M. 2010. *Brainstorm: The Flaws in the Science of Sex Differences*. Cambridge, MA: Harvard University Press.

Roughgarden, Joan. 2004. *Evolution's Rainbow: Diversity, Gender and Sexuality in Nature and People*. Berkeley: University of California Press.