

Resumen de *From Bacteria to Bach and Back: The Evolution of Minds*

Daniel Dennett

Las relaciones entre nuestras mentes (tal como las conocemos, precisamente por tenerlas) y nuestros cerebros (que podrían ser, por todo lo que sabemos de primera mano, órganos para refrigerar la sangre) están llenas de desconcertantes indicios y problemas y el tema conductor del libro es que se necesita entender cómo y por qué han evolucionado las mentes si se quiere entender lo que son ahora, y por qué tenemos tremendos poderes cognitivos y, a la vez, ilusiones cognitivas endémicas sobre esos problemas. Resulta que la introspección cartesiana, el tradicional “punto de vista de la primera persona”, no es la gran avenida que conduce al conocimiento, sino un persistente obstáculo para una captación más profunda de lo que sucede de forma efectiva en nuestras cabezas cuando pensamos, soñamos, sufrimos, esperamos o imaginamos. Antes de abordar un análisis de la experiencia consciente y de la naturaleza del *sujeto* de la experiencia, necesitamos valorar los tres mil millones y medio de historia de I+D –investigación y desarrollo– por medio de la selección natural, que ha dado forma a los circuitos, sistemas, herramientas de pensamiento y luchas intestinas que hacen todo el trabajo. Algunas de esas herramientas de pensamiento son productos del diseño inteligente –cálculo, estadísticas, microscopios, mapas– pero las innovaciones clave que hacen que las menes humanas sean mucho más poderosas que las mentes de otros animales son ellas mismas productos evolutivos de I+D, aunque de evolución *cultural*, no de evolución *genética*. Las palabras, de manera muy importante, pero también muchos otros modos de actuar y pensar que no pueden proferirse mediante el lenguaje, no fueron inventados por personas inteligentes, ni fueron adoptados y transmitidos porque ciertas personas fueran lo suficientemente inteligentes como para

darse cuenta de su utilidad. Se expandieron de la manera que lo hacen los virus, sin que los que los hospedan llegaran casi a darse cuenta y ciertamente sin que esos huéspedes los *entendieran* del modo en que hoy, retrospectivamente, podemos entenderlos.

Las ideas claves exigidas para entender esta afirmación son *razones de libre flotación* [*free-floating racionales*] y *competencia sin comprensión*. Tómese una simple bacteria; de hecho un pequeño agente prodigiosamente eficiente, complejo, que se protege a sí mismo, con muchos millones de partes que se mueven. Hay *razones* de *por qué* las partes tienen la estructura que tienen y llevan a cabo las tareas que realizan —razones de ingeniería, desde luego— pero ni las partes ni el todo de la bacteria necesita entender esas razones para beneficiarse de ellas. A lo largo de millones de años desde la aparición de la vida, la selección natural ha dado forma a una serie de micro-*gadgets* fundamentales —el ARN y después las secuencias de ADN, las proteínas motoras, los enzimas correctores de pruebas, las proteínas fotosensibles, los bulbos terminales sinápticos, los flagelos, ...— de la caja de herramientas que es compartida por todos los seres vivos y utilizada para dar forma a todos los diferentes órganos, inclinaciones y métodos de todos los agentes, desde las hormigas a los oligarcas, pasando por los robles.

Estas razones son de “flotación libre” porque *no están representadas en mente alguna* hasta que los inteligentes biólogos las descifran. Son razones que existen independientemente de ningún agente razonador. En el último siglo se han hecho grandes progresos, pero hay aún miríadas de razones de flotación libre que aguardan ser descubiertas, representadas, analizadas y confirmadas. Y estas dan cuenta de la competencia de los higos estranguladores (de los que no tenemos clave alguna sobre su agresiva explotación de los árboles que son sus compañeros), de los polluelos de cuco (de los que no sabemos lo que hacen), e incluso de los seres humanos que encuentran que determinadas cosas les producen asco (típicamente por buenas razones que sólo ahora han sido descubiertas). Esta es la “extraña inversión del razonamiento” de Darwin: la mente y la comprensión son productos recientes, no las fuentes de todo el diseño de la naturaleza.

No es ningún misterio el cómo este diseño se logró: la selección natural, sin ser consciente de ello, extrae valiosa información de los encuentros que todo ser viviente tiene con su entorno, buscando con ello unas organizaciones de las partes en funcionamiento más elegantemente ensambladas que, a su vez, permiten procesos aún más rápidos de exploración del espacio de diseño, hasta que, eventualmente, las condiciones estén

maduras para que una línea de descendencia (hasta ahora) genere un segundo proceso de selección natural, de *modos de hacer cosas* que se transmiten perceptual o socialmente, no a través de la línea germinal como instintos. Siguiendo a Dawkins (1976) llamo a esos modos *memes*. La iniciación del proceso de evolución cultural ha sido la transición reciente de mayor importancia [Szathmáry and Maynard Smith (1995)] y esto separa nuestra especie de todas las demás por lo que respecta a nuestra capacidad de comprensión.

Esta capacidad puede explicarse a su vez con la ayuda de otra extraña inversión del razonamiento debida, principalmente, a Alan Turing, que fue el primero en apreciar cómo un conjunto de interruptores (transductores y efectores) en última instancia mecánicos y carentes de mentalidad podrían organizarse en computadores aún más potentes, que —a diferencia de las personas que se denominaban “computadores” en la época de Turing— no necesitaban entender de aritmética para usarla sin cometer errores. Con aritmética y ramificación condicional, se podía realizar cualquier computación con tiempo y espacio de almacenaje suficientes. Los computadores, en el sentido moderno, proporcionan entonces vehículos de exploración para mirar al espacio de posibles arquitecturas cognitivas. Las neuronas no son muy parecidas a los circuitos o a los transistores, pero puede verse ahora claramente que su papel es computacional y, del mismo modo que hemos llegado a entender cómo las cosas de tamaño cotidiano pueden ser rojas sin que sus constituyentes últimos, átomos y moléculas, no sean en absoluto coloreados, podemos entender ahora cómo cosas de tamaño cotidiano como las personas pueden entender y ser conscientes sin que sus partes constituyentes comprendan o sean conscientes. La imagen manifiesta es una ilusión del usuario que ha evolucionado para capacitar un control más efectivo de cuerpos que contienen miles de millones de células, tanto por los mismos cuerpos como por otros. El que yo trate a otro como un agente consciente tiene que ver con el hecho de que yo saque ventaja de la ilusión del usuario de que el cuerpo que está enfrente de mí tiene un yo y una corriente de experiencia que comprende, y la misma actitud de usuario amistoso me ayuda a entender las circunstancias y las acciones del cuerpo que llamo “mío”. Una cascada de competencias sin comprensión puede generarla, dada la correcta organización. Y, finalmente, gracias a las miríadas de micro-actividades de esas piezas neurales que, en última instancia, son computacionales, con la comprensión llega la conciencia, la ilusión del usuario del cerebro de sus propias actividades, una ilusión perfectamente real que simplifica grandemente las complejas actividades que ocurren en el cerebro. ¿Quién o qué es el

“sujeto” de la ilusión, el “consumidor” (Millikan) de las *representaciones* internas? El controlador cuyas desenvueltas intervenciones mantienen al cuerpo lejos del peligro. Este controlador no es sino una organización de agencias neurales inconscientes, pero su capacidad de comunicarse con otros acerca de lo que está haciendo y por qué, le da una autoridad —no infalibilidad— de un “punto de vista de la primera persona”.

Ahora que podemos empezar a ver cómo construir mentes semejantes a las nuestras a partir de partes más elementales y sus operaciones, nos enfrentamos a una curiosa cuestión: ¿para qué es la comprensión? y ¿podríamos abandonar la comprensión como meta de la ciencia y disfrutar sólo de los frutos de las invenciones darwinianas (o darwinescas) tales como el aprendizaje mecánico y los algoritmos genéticos que están más allá de nuestra comprensión? Del mismo modo que la bacteria se beneficia de toda la brillante ingeniería que va en su estructura interna sin tener que entender nada de ella, quizás podamos proponernos ser beneficiarios de nuestros nuevos oráculos computacionales emergentes sin descartar la meta de explicar cómo todo esto nos sucede a nosotros y a los demás. Pero hay buenas razones —razones que podemos articular, evaluar y actuar sobre ellas— para mantener nuestra búsqueda de la comprensión.

Center for Cognitive Studies
Tufts University
115 Miner Hall
Medford, MA 02155 USA
E-mail: daniel.dennett@tufts.edu

REFERENCIAS BIBLIOGRÁFICAS

- DENNETT, D. (2017), *From Bacteria to Bach and Back: The Evolution of Minds*, London, Penguin
- DAWKINS, R. (1976), *The Selfish Gene*, Oxford, Oxford University Press. Versión castellana: *El gen egoísta*, Barcelona, Salvat editores.
- SZATHMÁRY E. and J. MAYNARD SMITH (1995), *The Major Transitions in Evolution*, Oxford, Oxford University Press.

Précis of *From Bacteria to Bach and Back: The Evolution of Minds*

Daniel Dennett

The relationship between our minds (as we know them just by having them) and our brains (which might, for all we know first-hand, be organs for cooling the blood) is full of baffling hints and problems, and the guiding theme of the book is that you need to understand how and why minds have evolved if you want to understand what they are now and why we have both tremendous cognitive powers and endemic cognitive illusions about those powers. It turns out that Cartesian introspection, the traditional “first-person point of view”, is not the royal road to understanding, but a persistent hindrance to a deeper grasp of what is actually going on in our heads when we think, dream, suffer, hope, imagine. Before we tackle an analysis of conscious experience, and the nature of the *subject* of experience, we need to appreciate the three-and-a-half-billion-year history of R&D — research and development — by natural selection, that has fashioned the circuits, systems, thinking tools, and internecine competitions that do all the work. Some of these thinking tools are themselves products of intelligent design — calculus, statistics, microscopes, maps — but the key innovations that make human minds so much more powerful than the minds of other animals are themselves products of evolutionary R&D, but of *cultural*, not *genetic*, evolution. Words, most importantly, but also many other culturally transmitted ways of acting and thinking that cannot be pronounced, were not invented by clever people, nor were they adopted and transmitted because people were clever enough to appreciate their utility. They spread the way viruses spread, almost unnoticed by their hosts and certainly not *understood* by their hosts the way we, today, can retrospectively understand them.

The key ideas required to understand this claim are *free-floating rationales* and *competence without comprehension*. Take a “simple” bacterium — in fact a breathtakingly efficient, complex, self-protective little agent with many millions of moving parts. There are *reasons why* the parts have the structure they do and perform the tasks they do — reasons of engineering, in effect — but neither the parts nor the whole bacterium needs to understand these reasons to benefit from them. Over millions of years since the dawn of life, natural selection shaped fundamental micro-gadgets — RNA and then DNA sequences, motor proteins, proof-reading enzymes, photosensitive proteins, synaptic endbulbs, flagella, ... — into the tool kit that is shared by all living things, and utilized to fashion all the different organs, inclinations and methods of all agents, from ants to oaks to oligarchs.

These reasons are “free-floating” because they are *not represented in any mind* whatever — until clever biologists figure them out. These are reasons that exist independently of any reasoners. Great progress has been made in the last century, but there are still myriads of free-floating rationales yet to be uncovered, represented, analyzed, and confirmed. And they account for the *competence* of strangler figs (which have no clue about their aggressive exploitation of their fellow trees) and cuckoo chicks (who know not what they do) and even human beings who find some things disgusting (typically for good reasons only now being discovered). This is Darwin’s “strange inversion of reasoning”: Mind, and Understanding, are recent products, not the sources of all the design in nature.

There is no mystery about how this design was accomplished: natural selection unwittingly extracts valuable information from the encounters that every living thing has with its environment, thereby homing in on ever more elegantly engineered, efficient organizations of the working parts, which in turn permit still swifter processes of exploration of Design Space, until eventually, the conditions were ripe for one lineage (so far) to generate a second process of natural selection, of *ways of doing things* that are transmitted perceptually or socially, not through the germ line as instincts. Following Dawkins (1976), I call these ways *memes*. The initiation of the process of cultural evolution was the most recent major transition [Szathmary and Maynard Smith (1995)], and it separates our species from all others in our capacity to understand.

This capacity in turn can be explained with the help of another strange inversion of reasoning, due, in the main, to Alan Turing, who first appreciated how an ultimately mechanical, mindless set of switches (simple transducers and effectors) could be organized into ever more powerful computers which — unlike the people who were called com-

puters in Turing's day — didn't need to understand arithmetic in order to do it flawlessly. With arithmetic and conditional branching, any computation can be performed, given enough time and storage space. Computers in the modern sense then provide the exploration vehicles for looking at the space of possible cognitive architectures. Neurons are not much like flip-flops or transistors, but their role can now clearly be seen to be computational. And, just as we have come to understand how everyday-sized things can be red without their ultimate constituents — atoms and molecules — being colored at all, we can understand how such everyday-sized things as persons can understand and be conscious without their constituent parts understanding or being conscious.

The manifest image is a user-illusion that has evolved to enable more effective control of bodies containing trillions of cells, both by those bodies themselves and by other bodies. My treating you as a conscious agent is a matter of my taking advantage of the user-illusion that the body in front of me has a self and a stream of experience which it understands, and the same user-friendly stance helps me understand the circumstances and actions of the body I call mine. A cascade of uncomprehending competences can thus deliver comprehension, given the right organization. And finally, thanks to the myriad micro-activities of these ultimately computational neural pieces, with comprehension comes consciousness, the brain's user-illusion of its own activities, a perfectly real illusion that hugely simplifies the complex activities occurring in the brain. Who or what is the "subject" of the illusion, the "consumer" (Millikan) of the inner *representations*? The controller whose fluent interventions keep the body out of harm's way. That controller is itself nothing but an organization of unconscious neural agencies, but its ability to communicate with others about what it is doing and why gives it the authority — not infallibility — of a "first person point of view."

Now that we can begin to see how to build minds like ours out of more elementary parts and their operations, we face a curious question: what is comprehension for, and might we abandon comprehension as a goal in science, and just enjoy the fruits of Darwinian (or Darwinesque) inventions such as machine learning and genetic algorithms, that are beyond our understanding? Just as bacteria benefit from all the brilliant engineering that went into their internal structure without having to understand any of it, perhaps we will settle for being beneficiaries of our newly emerging computational oracles while discarding the goal of explaining how it all happens to ourselves and others. But there are good

reasons — reasons we can articulate and evaluate and act on — for maintaining our quest for comprehension.

*Center for Cognitive Studies
Tufts University
115 Miner Hall
Medford, MA 02155 USA
E-mail: daniel.dennett@tufts.edu*

REFERENCES

- DENNETT, D. (2017), *From Bacteria to Bach and Back: The Evolution of Minds*, London, Penguin.
- DAWKINS, R. (1976), *The Selfish Gene*, Oxford, Oxford University Press.
- SZATHMÁRY E. and J. MAYNARD SMITH (1995), *The Major Transitions in Evolution*, Oxford, Oxford University Press.