Philosophical Psychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cphp20

Tracing the origins of consciousness

The Deep History of Ourselves: The Four-Billion-Year Story of How We Got Conscious Brains, by Joseph LeDoux, illustrated by Caio Sorrentino, New York, Viking, 2019, xviii+412 pp., USD $30, (Hardcover), ISBN 9780735223837

Jorge Morales

To cite this article: Jorge Morales (2023) Tracing the origins of consciousness, Philosophical Psychology, 36:4, 767-771, DOI: 10.1080/09515089.2023.2176746

To link to this article: https://doi.org/10.1080/09515089.2023.2176746

PHILOSOPHICAL PSYCHOLOGY
2023, VOL. 36, NO. 4, 767–771

Routledge
Taylor & Francis Group

BOOK REVIEW

Tracing the origins of consciousness

The Deep History of Ourselves: The Four-Billion-Year Story of How We Got Conscious Brains, by Joseph LeDoux, illustrated by Caio Sorrentino, New York, Viking, 2019, xviii+412 pp., USD $30, (Hardcover), ISBN 9780735223837

Darwin's theory is often illustrated with depictions of different finch beaks or with lined-up skeletons displaying subtle bone-structure changes throughout evolutionary history. In *The Deep History of Ourselves*, Joe LeDoux paints a fascinating picture in the reader's mind that contrasts not only organisms' shapes but also the subtle yet constant behavioral changes that govern life on Earth.

LeDoux does an extraordinary job showing that basic human behaviors can be traced all the way back to the origins of life. Old thermostats were round because they were controlled by dialing them to arrive at a target temperature. Modern, digital thermostats are often round too, not because their shape serves any function anymore, but as a vestige of an older era. Analogously, to understand the basic operations of human behavior as well as our complex cognition, emotions, and conscious awareness, the book argues that we need to take a dive into our deep history going back to the emergence of behavior and life themselves. In 66 short chapters neatly weaved together with this goal in mind, LeDoux uses the transition from unicellular to multicellular organisms, life outside the water, the appearance of animals, and the emergence of central nervous systems, to take the reader through the history of biology, evolution, neuroscience, and the acute minds that have put together the pieces of our origins.

LeDoux, whose prominent academic career has been focused on the neural basis of survival circuitries, emotions, and consciousness, claims that "behavior is not, as we commonly suppose, primarily a tool of the mind" (p. 11). Evolutionarily-relevant behavior predates the mind. The first living organism—possessing nothing that can meaningfully be called *mental*—is thought to have appeared 3.7 billion years ago, whereas neurons (let alone minds) appeared only approximately 600 million years ago. In this vein, he confesses: "my default position is that behavior is controlled nonconsciously until proven otherwise" (p. 328). The conscious feelings and emotions that we often take to be so definitional of our psyches could be unconnected from our most basic behaviors. A tempting mistake—*the great illusion*—is to think that because "humans consciously experience feelings when we engage in our own survival behaviors, we intuit that these feelings and the behaviors must be intrinsically related – that the feelings are the causes of the behaviors." (p. 3) LeDoux reserves conscious experiences and emotions to hierarchical higher-order processes that involve the self, memory systems, and the brain's ability to make predictions. The rest is just behavior of the kind inherited from our long-dead, simple-minded (and even "non-minded") ancestors.

After a 100-page erudite journey in evolutionary biology, the rest of the book focuses, first, on the emergence of neurons, nervous systems, and the mammalian brain. Then, it zooms on the mind: internal representations that sustain flexible

behavior and deliberation as well as perceptual and mnemonic systems. In one of the most exciting parts of the book, LeDoux the anatomist and cognitive neuroscientist, describes in painstaking detail how the prefrontal cortex became the hub of complex cognitive processing, memory, and prediction. It is in this context that consciousness makes its appearance (both in the book and in evolutionary history).

LeDoux, now wearing his theorist and philosopher caps, proposes a higher-order theory (HOT) of consciousness. Like in other HOTs, awareness emerges when a higher-order state (e.g., a thought) represents the subject as being in a targeted first-order state (typically, a sensory state). A huge advantage of HOTs (although not exclusive of them) is that they naturally allow for the separation of information processing (first-order states) and consciousness. The main innovation of LeDoux's multistate hierarchical model of consciousness is that the network is multilayered. The node at the top of the network (speculatively housed in the frontal polar cortex; p. 309) receives mostly multimodal (i.e., conceptual and mnemonic) inputs as well as inputs from other prefrontal areas. These, in turn, are higher-order with respect to lower-order, unimodal sensory inputs. This multimodal intermediate layer contains the predictions and schemas required for conscious experiences to efficiently emerge in everyday, complex situations. An important outcome of the multistate hierarchical arrangement is that the states processed in the prefrontal cortex are "diverse and involve partly redundant representations with different degrees of abstraction that reflects their origin in multiple processing streams (sensory, memory, conceptual)" (p. 308). A consequence of this is that the neural focus of awareness may change as an episode develops over time. This is one of the most enticing hypotheses in the book—one that has not been sufficiently tested and that might help harmonize discrepant findings in the literature. This feature of the theory prevents a common argument from opponents of HOTs who have critically asked why lesions in the prefrontal cortex do not always result in complete abolition of awareness (e.g., Boly et al., 2017; Kozuch, 2013; Kozuch, 2022; but see Michel, 2022; Odegaard et al., 2017). LeDoux's answer: the system is dynamic, distributed, and it has many backdoors and backups.

While this kind of higher-order architecture can explain why we are aware of objects in the world, it is insufficient for explaining self-awareness (awareness of ourselves as subjects)—a capacity LeDoux reserves to human beings alone (p. 310). The ability to use autobiographical memory in thinking about one's self (also known as autonoetic consciousness) requires a further self-schema, which is fundamental in experiencing (human) emotions (a whole different category of conscious experiences). "No self, no fear" is LeDoux's motto. Survival circuitry states underlie behavioral and physiological avoidance responses in the case of fear (e.g., sweating, freezing, and fleeing), but they are not the emotion of fear. Naturally, nonconscious representations (schemas) that help make sense of the situation in which we find ourselves are important for experiencing emotions. Self- and emotion-schemas are of particular notability. But, ultimately, human emotions are "cognitively assembled" (p.351). Emotions require the agent conceptualizing themselves in a particular situation (Wilson-Mendenhall et al., 2011). Awareness of danger is not the same as knowing that *you* are in danger. Circling back to the deep history *leitmotiv*, LeDoux ends the book reflecting on how survival behavior is deep, but emotions are shallow—in other words: emotions are recent, uniquely human experiences.

§

LeDoux's "basal cognition" approach (Levin, 2021; see Colaço, 2022; Figdor, 2018) tracing a continuum between homeostatic mechanisms in ancient cells as well as complex learning and goal-directed behavior is presented in a very attractive way as

one flips the pages. The transitions are smooth, and the narrative is well crafted. In what can only be described as a book-length evolutionary sorites, it is easy to see how a particular organism solves problems the way they do based on their proximal ancestor's solution to a similar problem and the same applies for their descendants and their descendants and so on. However, one of the main contentions of the book is that despite the *prima facie* lack of intuitiveness of linking the extremes of the chain (as in any other sorites paradox) these too can be tied up with some ingenuity. Nonetheless, the apparent paradox may not be so easily solved.

While it might be true that "behavior, learning, and memory don't actually require a nervous system" (p. 2), this does not mean that the solutions found by creatures with nervous systems can be traced back to creatures without them. (Also note that it is contentious that *learning* and *memory* proper can take place without nervous systems and without representations). After all, it is not the case that the *same* mechanisms that produced some behavior in ancient organisms are preserved *in the same* way in modern animals with just more features added. Although LeDoux sometimes writes in that way: "only by examining the natural history of life on Earth can we appreciate the unique features that were added [sic] to organisms incrementally over the course of evolution" (pp. 11–12). Organisms' features are transformed, not merely added.

It is perhaps this strong commitment to a deep history that leads to a couple of what may be among the most controversial claims in the book: (i) that most behavior is controlled unconsciously and (ii) that with the exception of mammals and birds, animals lack cognitive sophistication (and all of them lack the ability to experience emotions).

LeDoux writes different versions of the following throughout the book: "I assume brains were nonconscious long before they were conscious. So my default position is that behavior is controlled nonconsciously until proven otherwise" (p. 328). The second claim need not follow from the first, even if it is definitely true that brains predate consciousness. That x can be done without y does not mean that x is done without y in all contexts. I could eat soup without a spoon, but that does not entail I normally do. That primitive, unconscious brains could display complex behavior does not entail that consciousness is not normally used to perform similar behaviors by conscious creatures. Other scientists and philosophers have also wondered about the exact degree of conscious control we have over our behavior, and the answer is far from certain (Koizumi et al., 2015; Robinson et al., 2015; Rosenthal, 2008). Nevertheless, it is unclear why the science of consciousness would supposedly "advance much more smoothly" (p. 50) if we were to commit to LeDoux's default position. It is undeniably true that many of our (relatively basic) behaviors are controlled automatically. However, skepticism works both ways: our strong introspective evidence that conscious experiences are causally relevant for many of our behaviors should also give us pause thinking that they can take place unconsciously until proven otherwise.

Throughout the book, the decision to exclude some species from the "cognitive realm" is problematic. LeDoux defines cognition (pretty uncontroversially) as the ability to "form representations and use them to guide behavior" (p. 34). But then, he writes that "basic cognitive capacities, like outcome-dependent instrumental responses [and flexible goal-directed learning] are present in mammals and birds, but have not been demonstrated in other species" (p. 34; see p. 219). It is surprising that bees and octopuses, to list just two examples, are left out from this select group.

Bees have complex navigational systems thought to require representational maps (Gallistel, 2011). They can be trained through social learning to deploy unusual behaviors as they learn to expect reward (Loukola et al., 2017; see Gil et al., 2007; Niv

et al., 2002). Octopuses have also demonstrated cognitive flexibility, object permanence, and spatial navigation consistent with a sophisticated cognitive apparatus (Godfrey-Smith, 2016; Schnell & Clayton, 2019).

With these blind spots in mind, one cannot help but wonder whether a more expansive view of consciousness is possible too – perhaps not for bees and octopuses, although it is certainly possible, but at least for other mammals, including other primates. LeDoux makes a point of showing that several structures and functional connectivity profiles of the prefrontal cortex are unique to humans, suggesting that this is perhaps why we are the only creatures with self-awareness. But a lesson from *The Deep History of Ourselves* is that "overall similarities outweigh differences between the brain of new versus ancestral species. Human brains are no exception" (p. 259). There is no denying that animal consciousness and self-awareness (including emotions) may be different from ours (it would be strange if they were identical). However, just as there is behavior and perhaps even memory and learning without a nervous system, it is perhaps possible to have some kind of self-awareness in animals without a human prefrontal cortex.

The claim that animals can be self-aware is, of course, highly controversial. I am, in fact, uncommitted in the case of most species. But as with the role of consciousness discussed above, here too skepticism should work both ways: given our evolutionary proximity to other species and the deep history of behavior and cognition, it would not be strange that non-human forms of self-awareness are indeed present in at least some other species. Given the paucity of definitive evidence either way, it could turn out that emotions are not so shallow after all.

These points of contention do not take away from the immensely rewarding experience that is reading *The Deep History of Ourselves*. The book gifts the reader – both in and outside academia – with a mix of encyclopedic knowledge and insightful reflections on the origins of behavior and the human mind, all accompanied with gorgeous illustrations. Moreover, with a theoretically sound and neuroscientifically supported theory of consciousness, LeDoux offers a serious alternative over other available options.

References

Boly, M., Massimini, M., Tsuchiya, N., Postle, B. R., Koch, C., & Tononi, G. (2017). Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? Clinical and neuroimaging evidence. *Journal of Neuroscience*, *37*(40), 9603–9613. https://doi.org/10.1523/jneurosci.3218-16.2017

Colaço, D. (2022). Why studying plant cognition is valuable, even if plants aren't cognitive. *Synthese*, *200*(6), 453. https://doi.org/10.1007/s11229-022-03869-7

Figdor, C. (2018). *Pieces of mind: The proper domain of psychological predicates*. Oxford University Press.

Gallistel, C. R. (2011). Prelinguistic thought. *Language Learning and Development*, *7*(4), 253–262. https://doi.org/10.1080/15475441.2011.578548

Gil, M., Marco, R. J. D., & Menzel, R. (2007). Learning reward expectations in honeybees. *Learning & Memory*, *14*(7), 491–496. https://doi.org/10.1101/lm.618907

Godfrey-Smith, P. (2016). *Other minds: The octopus, the sea, and the deep origins of Consciousness*. Farrar, Straus and Giroux.

Koizumi, A., Maniscalco, B., & Lau, H. (2015). Does perceptual confidence facilitate cognitive control? *Attention, Perception, & Psychophysics*, *77*(4), 1295–1306. https://doi.org/10.3758/s13414-015-0843-3

Kozuch, B. (2013). Prefrontal lesion evidence against higher-order theories of consciousness. *Philosophical Studies*, *167*(3), 721–746. https://doi.org/10.1007/s11098-013-0123-9

Kozuch, B. (2022). Underwhelming force: Evaluating the neuropsychological evidence for higher-order theories of consciousness. *Mind & Language*, *37*(5), 790–813. https://doi.org/10.1111/mila.12363

Levin, M. (2021). Life, death, and self: Fundamental questions of primitive cognition viewed through the lens of body plasticity and synthetic organisms. *Biochemical and Biophysical Research Communications*, *564*, 114–133. https://doi.org/10.1016/j.bbrc.2020.10.077

Loukola, O. J., Solvi, C., Coscos, L., & Chittka, L. (2017). Bumblebees show cognitive flexibility by improving on an observed complex behavior. *Science*, *355*(6327), 833–836. https://doi.org/10.1126/science.aag2360

Michel, M. (2022). Conscious perception and the prefrontal cortex a review. *Journal of Consciousness Studies*, *29*(7), 115–157. https://doi.org/10.53765/20512201.29.7.115

Niv, Y., Joel, D., Meilijson, I., & Ruppin, E. (2002). Evolution of reinforcement learning in foraging bees: A simple explanation for risk averse behavior. *Neurocomputing*, *44*, 951–956. https://doi.org/10.1016/s0925-2312(02)00496-4

Odegaard, B., Knight, R. T., & Lau, H. (2017). Should a few null findings falsify prefrontal theories of conscious perception? *Journal of Neuroscience*, *37*(40), 9593–9602. https://doi.org/10.1523/JNEUROSCI.3217-16.2017

Robinson, Z., Maley, C. J., & Piccinini, G. (2015). Is consciousness a spandrel? *Journal of the American Philosophical Association*, *1*(2), 365–383. https://doi.org/10.1017/apa.2014.10

Rosenthal, D. (2008). Consciousness and its function. *Neuropsychologia*, *46*(3), 829–840. https://doi.org/10.1016/j.neuropsychologia.2007.11.012

Schnell, A. K., & Clayton, N. S. (2019). Cephalopod cognition. *Current Biology*, *29*(15), R726–732. https://doi.org/10.1016/j.cub.2019.06.049

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., & Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia*, *49*(5), 1105–1127. https://doi.org/10.1016/j.neuropsychologia.2010.12.032

Jorge Morales
Department of Philosophy, Northeastern University, Huntington Ave, Boston, MA
Department of Psychology, Northeastern University, Huntington Ave, Boston, MA
j.morales@northeastern.edu http://orcid.org/0000-0001-5825-4909

https://doi.org/10.1080/09515089.2023.2176746