Review of: "Growing Confidence and Remaining Uncertainty About Animal Consciousness"

Pieter de Vries[1]

1 University of Groningen

Potential competing interests: No potential competing interests to declare.

Irwin 2024, Growing Confidence and Remaining Uncertainty About Animal Consciousness

Goal of the paper

The well-written and highly readable paper gives an overview of the various studies on human and animal consciousness covering the last fifty years. Its goal is to determine for which topics substantial scientific agreement exists and which ones are still open to considerable dispute.

The number of studies and their relevance to the topic is impressive, especially if we consider the large variety of involved disciplines and the many aspects of consciousness.

Comments

The author's conclusion about the relatively large agreement existing for definitional and ontological aspects of consciousness, as well as about the role of qualia, mental causation, and the self, ties in with the hypothesis that consciousness is – scientifically speaking, and despite its intangible nature – a normal property. Like many other properties in science, or in daily life, it can be seen as the product of several or possibly many interacting elements. As such, it is a new property since none of these elements does have this property, not even to a lesser degree. Examples are abundant. In chemistry, neither oxygen nor hydrogen has the property of 'wetness' characteristic of their chemical composition in water, or, in daily life, none of the parts of a chair has – in isolation - the property of 'comfort of sitting' of a chair that is assembled of these parts.

The study of the 'consciousness-as-a-normal-scientific-property' hypothesis implies a systems approach to the study of the nervous system. Given the historical backgrounds in the paper, I suggest that the incorporation of the contributions of one of the pioneers of this approach, Sommerhoff (1994; 1996), may be worth considering. He emphasised the importance of internal representations in the development of consciousness in systems, such as organisms, men, and machines. By means of a specification of internal representations of a system, a more satisfactory answer can be given to the question of what is meant by the complexity of the neural, and potentially other, mechanisms underlying consciousness. In the paper, this complexity topic pops up several times.

Taking a larger perspective on the notion of representation, Dalenoort (1990) focuses on the relative autonomy of different levels of explanation. This brings us to the issues about consciousness that, according to the reviewed paper, are still in

dispute.

These remaining uncertain issues, as they are referred to by the author, are: the substrate of consciousness, the gap between mechanism and phenomenon, and a 'neural monitor' required for consciousness.

As currently formulated in the paper, these three issues can be reduced to two, or maybe even to one. So, apparently, there may be less uncertainty. In my opinion, this should be done because the discussion of the first issue, the substrate, is in fact about three general models or architectures for cognitive brain functioning and underlying neural processes: the dynamic thalamocortical core, the global workspace, and higher-order thought. If these models are to be complete, they should also include hypotheses about the so-called 'neural monitor', i.e., the third uncertain issue. In line with Sommerhoff (1996), these hypotheses may refer to properties of neural, internal representations of a(n) (human) organism. A generalisation of Hebb's notion of a cell-assembly, quoted in the paper, then offers a starting point for further study, cf. De Vries (2004, 2019; 2008), especially when combined with solutions for the various forms of the binding problem discussed by Feldman (2013).

In relation to the second uncertain issue, the gap between mechanism and phenomenon, it may be useful to consider the framework of relatively autonomous levels of explanation, cf. Dalenoort (1998). According to this framework, a gap is always present if – in the description of a phenomenon - we go from one level of explanation to a higher one. This is shown by the examples given above. Following this framework, what we can do when we explain a phenomenon is specify the conditions at one level that are necessary and sufficient to produce a certain phenomenon at a higher level. This is then a form of non-eliminative reductionism. Alternatively, we can look for the functions of a phenomenon and try to specify possible conditions for its existence and development at a lower level. We then have a form of non-eliminative functionalism. I use the known term 'non-eliminative' to indicate that both approaches are no longer restricted to one level of explanation, as they once used to be, cf. Fodor and Pylyshyn (1988) and Churchland (1989).

Specifically in relation to consciousness, the work of Velmans (2000) may be worth considering in relation to the reductionism-functionalism discussion. In line with the arguments above, he concludes that "mind and brain are two sides of the same coin", implying that you *are* not your brain. Likewise, Velmans' maxim holds for non-eliminative reductionism and functionalism. They do not exclude each other and are both necessary for gaining more insight into a phenomenon.

These are my suggestions for additions to an already very good paper.

References

Churchland, P. S. (1989). *Neurophilosophy, Toward a unified science of mind/brain* MIT Press.

Dalenoort, G. J. (1990). Towards a general theory of representation. *Psychological Research*, *52*(2), 229–237.

Dalenoort, G. J., & de Vries, P. H. (1998). Understanding in terms of correspondences between levels of description. In C. Taddei-Ferretti & C. Musio (Eds.), *Downward processes in the perception representation mechanisms* (1–Book, Section, pp. 497–503). World Scientific.

de Vries, P. H. (2004). Effects of binding in the identification of objects. *Psychological Research*, *1*(2), 41–66.

de Vries, P. H. (2019). Conditions for Cognitive Self-Organisation Implied by Visual-Word Processing. *Manuscript Submitted for Publication*, *Journal Article*.

de Vries, P. H., & van Slochteren, K. R. (2008). The nature of the memory trace and its neurocomputational implications. *Journal of Computational Neuroscience*, *25*(Journal Article), 188–202.

Feldman, J. (2013). The neural binding problem(s). *Cognitive Neurodynamics*, *7*(1), 1–11. https://doi.org/10.1007/s11571-012-9219-8

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(Journal Article), 3–71.

Sommerhoff, G. (1996). Consciousness explained as an internal integrating system. *Journal of Consciousness Studies*, *3*(2), 139–157.

Sommerhoff, G., & MacDorman, K. (1994). An account of consciousness in physical and functional terms—A target for research in the neurosciences. *Integrative Physiological and Behavioral Science*, *29*(2), 151–181. https://doi.org/10.1007/BF02691012

Velmans, M. (2000). *Understanding consciousness*. Routledge.