

Letter

Do You Need to Be
Conscious to Learn
to Be Conscious?

Axel Cleeremans,^{1,*} Dalila Achoui,¹
Arnaud Beauny,¹ Lars Keuninckx,²
Jean-Remy Martin,¹
Santiago Muñoz-Moldes,³
Laurène Vuillaume,¹ and
Adélaïde de Heering¹

Overgaard and Kirkeby-Hinrup conclude their comment [1] on our Opinion article [2] by asking, 'Does SOMA entail that animals, children and many adults are not conscious?' This question is indeed essential, and our answer is a clear: 'We do not know, nor does anyone else.' But we want to state right away that we *believe* they are. In this respect, Overgaard and Kirkeby-Hinrup's core critique is misguided. We need to carefully distinguish between ethical and scientific considerations. Scientific claims are based on what we think is the case, whereas ethical considerations are based on what we think ought to be the case. Thus, one might think that infants lack consciousness yet choose to act toward them as though they were. Hinduism considers that all living things are sacred and should not be harmed (*ahimsa*). Jainists go further and will wear masks to avoid accidentally breathing in insects. Such beliefs and practices are independent of our scientific knowledge. Although one may wish that our ethics be informed by science, the former is independent of the latter and should in no way be subsidiary to it. Appealing to ethical considerations to rebut scientific claims is rhetorically unsound.

What are the scientific claims that Overgaard and Kirkeby-Hinrup question, then?

First is the critique that the self-organizing metarepresentational account (SOMA)

requires too rich a conceptual apparatus to subtend consciousness. This is a well-known critique of higher-order approaches to consciousness, and it has been addressed to some extent by Gennaro [3] and by Brown *et al.* [4]. It is, in our view, based on a misunderstanding of what is actually required for Higher-Order Thought Theory (HOTT)'s core mechanism of rerepresentation to be effective. SOMA is precisely an attempt to deflate the requirement from canonical HOTT that 'thoughts' of any kind need be involved.

Second, Overgaard and Kirkeby-Hinrup question the idea that learning is instrumental to phenomenal awareness. Learning unsurprisingly shapes conscious experience, they say, but for the shaping to take place, there has to be phenomenal experience in the first place. Here, we argue again that there is a difference between sensitivity and awareness. Many living things, from plants to infants and even machines, can be characterized as being sensitive. In that sense, they have protofeelings of sorts – states that appropriately point to certain conditions and that are associated, minimally, with action tendencies and physiological responses. Crucially, however, such sensitivity need not be accompanied by conscious experience. Thus, although *something* senses, the agent in which that something is located does not feel it. The difference between sensitivity and awareness is in fact the core distinction that any theory of consciousness should aim to account for. Sensitivity is a background condition for consciousness, but it is not equivalent to it: Consciousness involves additional mechanisms through which the agent itself becomes sensitive to its own sensitivity. SOMA proposes that such mechanisms are the product of processes of learning and plasticity through which agents develop systems of metarepresentation that redescribe and qualify their own representational space.

Third, taken in its most extreme formulation, our proposal, as Overgaard and Kirkeby-Hinrup rightly claim, is that phenomenal consciousness depends on self-consciousness, and that self-consciousness in turn depends on Theory of Mind (ToM). SOMA, congruently with Carruthers's perspective [5], radically inverts the causal flow typically assumed to unfold over development, whereby one assumes that we begin with raw phenomenal experience, which is then augmented by an emerging awareness of one's existence as an agent (self-awareness), only to be followed much later by the ability to understand that other agents likewise have conscious mental states (ToM). Although the traditional perspective leaves phenomenal experience completely unexplained, SOMA's burden is to answer the question of *when* phenomenal awareness emerges. There is undoubtedly a 'before' and an 'after.' The question of when the transition occurs over the course of development remains, for now at least, very much in the realm of speculation, as is the thorny question of the extent to which this transition is best described as graded or dichotomous. But there is now substantial evidence that the relevant transitions appear much earlier than we thought (Box 1), perhaps even as early as during fetal life, and that they unfold gradually. Thus, we purport that the loopy relationships between phenomenal awareness, self-awareness, and ToM are set in motion very early on.

In conclusion, matter that thinks also thinks about itself. SOMA views consciousness as resulting from a largely self-organizing, socially driven predictive process through which agents gradually come to be sensitive to the existence, structure, and qualities of their own internal states in increasingly sophisticated ways that ultimately result in the construction of self, that is, of representations of one's own inner workings that are tuned toward enabling flexible control over action. *When* the conscious mind bootstraps itself into existence is a fascinating,

Box 1. Empirical Evidence for Early Competence

Just as animals and plants are getting smarter every year (they do not, but our detection methods have become more sensitive), there is now substantial evidence that infants are capable of much more than we assumed was the case even just 10 years ago. Infants as young as 5 months old are capable of implicit theory of mind – that is, the ability to correctly predict others' actions based on their goals [6]. Thus, although explicit ToM may only develop much later, even young infants are capable of correctly inferring other people's mental states in specific contexts. Likewise, just as is the case for animals (see [7] for a recent special issue), preverbal infants as young as 12 months old are capable of limited forms of metacognition [8]. Furthermore, Kouider *et al.* [9] showed how neural markers associated with consciousness are already present, though in a delayed form, in infants. There is also evidence that fetuses have social interactions, as an *in utero* study of the movement kinematics of twins has intriguingly suggested [10]. Finally, self-awareness itself is far from a monolithic concept and likely develops gradually in different directions [11,12]. These findings all suggest that core cognitive abilities come into existence much earlier than typically assumed.

gruesomely complex question, but every passing year brings more evidence that it occurs earlier and more gradually than we thought. SOMA does not require us to think of infants, animals, and adult persons with a ToM deficit as lacking consciousness, though there is little doubt that their experience is probably very different from that of typical human adults. But if consciousness begins with modeling other agents' mental states, as a radical interpretation of SOMA suggests, then we all bear responsibility in being the very best conscious agents we can.

Acknowledgments

We thank Matthias Michel for useful suggestions.

¹Consciousness, Cognition & Computation Group (CO3), Center for Research in Cognition and Neuroscience (CRCN), ULB Neuroscience Institute (UNI), Université libre de Bruxelles, B1050 Brussels, Belgium

²BeloT Neuromorphic Team, Interuniversity Microelectronics Centre (IMEC), Leuven, Belgium

³Cambridge Consciousness and Cognition Lab, Department of Psychology, University of Cambridge, Cambridge, UK

*Correspondence:
axcleer@ulb.ac.be (A. Cleeremans).
<https://doi.org/10.1016/j.tics.2020.10.002>

© 2020 Published by Elsevier Ltd.

References

- Overgaard, M. and Kirkeby-Hinrup, A. (2020) Is learning the cognitive basis of consciousness? The moral implications of SOMA. *Trends Cogn. Sci.* Published online October 6, 2020. [https://www.cell.com/trends/cognitive-sciences/fulltext/S1364-6613\(20\)30212-6](https://www.cell.com/trends/cognitive-sciences/fulltext/S1364-6613(20)30212-6)
- Cleeremans, A. *et al.* (2020) Learning to be conscious. *Trends Cogn. Sci.* 24, 112–123
- Gennaro, R.J. (2012) *The Consciousness Paradox*, MIT Press
- Brown, R. *et al.* (2019) Understanding the higher-order approach to consciousness. *Trends Cogn. Sci.* 23, 754–768
- Carruthers, P. (2009) How we know our own minds: the relationship between mindreading and metacognition. *Behav. Brain Sci.* 32, 121–138
- Sobian, B. *et al.* (2020) How does children's Theory of Mind become explicit? A review of longitudinal findings. *Child Dev. Perspect.* 14, 171–177
- Beran, M.J. (2019) Animal metacognition: a decade of progress, problems, and the development of new prospects. *Anim. Behav. Cogn.* 6, 223–229
- Goupil, S. and Kouider, S. (2016) Behavioural and neural indices of metacognitive sensitivity in preverbal infants. *Curr. Biol.* 26, 3038–3045
- Kouider, S. *et al.* (2013) A neural marker of perceptual consciousness in infants. *Science* 340, 376–380
- Castiello, U. *et al.* (2010) Wired to be social: the ontogeny of human interaction. *PLoS One* 5, e13199
- Rochat, P. (2003) Five levels of self-awareness as they unfold early in life. *Conscious. Cogn.* 12, 717–731
- Legrain, L. *et al.* (2010) Distinguishing three levels in explicit self-awareness. *Conscious. Cogn.* 20, 578–585