

The grand challenge of consciousness

Article (Published Version)

Seth, Anil K (2010) The grand challenge of consciousness. *Frontiers in Psychology*, 1. ISSN 1664-1078

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/56501/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.



The grand challenge of consciousness

Anil K. Seth*

Sackler Centre for Consciousness Science, School of Informatics, University of Sussex, Brighton, UK

*Correspondence: a.k.seth@sussex.ac.uk

Consciousness is at once the most familiar and the most mysterious aspect of our existence. Conscious experiences define our lives, but the subjective, private, and qualitative nature of these experiences seems to resist scientific inquiry. For much of the twentieth century, consciousness research remained the exclusive preserve of philosophy, whose practitioners continue to grapple with the so-called ‘hard problem’ of why physical or physiological processes should give rise to conscious experiences *at all*. Yet people have wondered about consciousness since they wondered about anything, and advances in our comprehension have been slow in coming.

Over the last two decades much has changed (Baars et al., 2003; Metzinger, 2003; Banks, 2009; Bayne et al., 2009). Alongside philosophical discourse a new science of consciousness has taken shape which integrates experimental and theoretical work across many fields including neuroscience, psychology, cognitive science, artificial intelligence, computer science, neurology, and psychiatry. Developing a naturalized account of the rich experiential tapestry of consciousness is now recognized as a major objective for twenty-first century science. Perhaps the key factor in the transition to scientific legitimacy was the realization that it may not be necessary to explain *why* consciousness exists in order to begin to unravel the physical and biological mechanisms that underlie its various properties. After all, physicists have laid bare many mysteries of the universe without accounting for the brute fact of its existence.

As this realization has taken hold, the excitement of consciousness research has become increasingly pervasive. I believe that this excitement is justified. Though progress may at times seem slow, by examining core questions in consciousness science we stand to learn a great deal about ourselves and our place in nature. Here, I offer a personal perspective on some of these core questions alongside some broader considerations raised by consciousness research.

A first core challenge is to specify the necessary and sufficient processes that underpin normal human consciousness (Edelman, 2003; Koch and Tononi, 2008). We already know that we can do without large parts of the brain (e.g., the cerebellum) without appreciable degradation of consciousness. Bodily and environmental interactions also seem optional, at least for consciousness *at a given time*, as shown by dreaming and conditions such as the locked-in state. Current consensus favors the thalamocortical system as the seat of the relevant neural machinery, however it is unknown which (if any) components of this system are critical, or whether it is the dynamical activity patterns that flow across its neurons and synapses that matter most. Key to addressing this challenge will be to move beyond looking for correlations between (conscious) phenomenal properties and neural properties, towards identifying ‘explanatory correlates’ that actually account for phenomenal properties (e.g., the simultaneously integrated and differentiated nature of experience) in terms of corresponding neural properties (Seth, 2009). Work in this direction, which will require both theoretical and experimental innovations, may help us answer probing questions such as why experiences of color are qualitatively different from experiences of sounds, or of odors.

Consciousness is not a unitary phenomenon. We can distinguish conscious *level* (a scale from coma or brain-death to fully vivid conscious awareness) and conscious *content* (the components of each conscious experience). Conscious contents themselves differentiate into (at least) multimodal sensory contents related to the world; experiences of selfhood, volition, and agency; and affective and somatic perceptions. We now need to understand the extent to which these components are separable, what their respective neural underpinnings may be, and how eventually they are integrated into the seamless flow of our normal conscious lives.

A related challenge lies in identifying the function (or functions) of consciousness. Setting aside the red herrings of

epiphenomenalism (the notion that consciousness has no function) and ‘conscious inessentialism’ (the idea that all cognitive and behavioral functions can in principle be carried out in the absence of consciousness) allows us to consider a range of possible functions for the different components of consciousness. Current candidates include supporting voluntary and/or rational action; enabling flexible and integrated responses to richly structured environments; acquiring new skills; correcting errors in perception and action; simulating potential threats (in dreams), and enabling effective social cognition. All of these remain controversial but most can be experimentally interrogated. An important avenue for further progress will be to better understand the capabilities and limitations of unconscious (implicit) processes, and the interactions between such processes and explicit, conscious states.

A mature science of consciousness requires effective means for measurement of conscious content and conscious state, both for mapping experimental evidence to theory and for designing perspicuous experiments. Measures of consciousness can be objective (e.g., behavioral responses, measured brain signals) or subjective (e.g., introspective reports, confidence ratings). While it is unlikely that any single measure will prevail, recent advances have involved combining in single experiments multiple measures of both kinds. More generally, solving the measurement problem will require distinguishing between neural mechanisms giving rise to consciousness *per se* from mechanisms that enable its subjective report: this distinction is sometimes referred to in terms of ‘phenomenal’ *versus* ‘access’ consciousness and it may also relate to the interaction between consciousness and self-hood.

Another fundamental challenge is to understand the nature of disorders of consciousness, as they affect both state (e.g., coma, the vegetative and minimally consciousness states, and epileptic absence seizures) and content (e.g., psychiatric

disorders, focal brain lesions leading to neglect and agnosia, and certain degenerative neurological illnesses). Work in this area is progressing particularly rapidly (Owen et al., 2009). New neuroimaging and psychophysical methods have allowed detection of residual consciousness in patients previously diagnosed as vegetative, allowing more effective prognosis and palliative care and in some cases restoring communication. New perspectives are also emerging in psychiatry. For example, schizophrenia has been related to deficiencies in fine-grained predictions of the consequences of self-generated actions, and 'depersonalization disorder' – a condition in which the world, and/or the self, loses its phenomenal 'reality' – may be associated with deficient integration of exteroceptive signals with interoceptive, autonomic signals. Quite apart from the obvious clinical significance of this research, disorders of consciousness will provide an increasingly important window onto the neural and cognitive mechanisms of unimpaired consciousness.

We may also look ahead to three challenges that might remain on the horizon for some time to come. Responses to these challenges, when they do arise, may however hold the greatest potential for transforming our understanding of our place in nature.

First, which non-human animals are conscious, and when did consciousness arise in evolution? (Edelman and Seth, 2009). Already many of us tacitly ascribe consciousness to primates and many other mammals, though ascription of full-fledged self-consciousness remains more controversial. Moving beyond mammals, perhaps a strong case can be made for birds (especially corvids and parrots) and there is tantalizing evidence of behavior consistent with consciousness in some cephalopods (e.g., *Octopus vulgaris*) though knowledge of corresponding neural properties is sorely lacking. But where do we draw the line? Can a case be made for consciousness in bees, or in fruit flies? And what would it take to ascribe a conscious 'I' to any non-human animal?

Consciousness and conscious self-hood must have beginnings in ontogeny as well as in phylogeny. The second challenge is therefore to identify when these properties arise

during human development. While this challenge may appear more tractable than that of identifying animal consciousness, it is not certainly so and potential answers carry heavy political and social implications. For example, how would society react if it were shown that babies remained unconscious until several months following birth?

Third, is it possible to devise a conscious artifact? If consciousness is assumed to depend on the laws of physics, chemistry, and biology, then however imperfectly these laws are currently known the answer must be yes. What is not known is what kind of artifact will be adequate. Contemporary computational and robotic models of consciousness may best be seen as simulations rather than instantiations, in the same sense that a computational model of a hurricane is not itself windy. However, as models of consciousness progressively build in constraints that counter theoretical objections and cater to empirical mismatches that arise from previous attempts, so the new models may tend towards instantiation. Of course, it is presently impossible to know whether a model that is sufficiently rich to account causally for all properties of consciousness will be implementable in computers or robots, or whether it will require implementation in neural or some other material.

The new knowledge gained as we unravel the mechanisms of consciousness will have substantial practical, ethical, moral and even legal consequences. Practically, we may witness a raft of new clinical interventions as well as new methods for manipulating conscious experiences in healthy subjects. Ethically and morally we may confront new dilemmas relating to clinical decision making and to the treatment of non-human animals. Legally, we may face complex issues involving assignment of responsibility, arising from new insights into the power of unconscious processes and into the neural mechanisms underlying volition and agency. This is of course an incomplete and labile list: its constitution will change as our understanding develops and deepens.

In conclusion, consider for a moment the pernicious perspective of 'mysterianism', the notion that a naturalized account of consciousness may exist but is 'cognitively

closed' to us in the same way that an appreciation of quantum mechanics is inaccessible to frogs. While the possibility cannot be ruled out, accepting it achieves nothing and the history of human intellectual ingenuity suggests that there is plenty more we can know that we cannot yet conceive of. The grand challenge of consciousness rests on this hope. Its pursuit is distinguished by combining transformative questions about the human condition with a tractable and largely incremental programme of experimental and theoretical research. There is every reason to be optimistic.

ACKNOWLEDGMENTS

Anil K. Seth is supported by EPSRC Leadership Fellowship EP/G007543/1 and by the Dr. Mortimer and Theresa Sackler Foundation.

REFERENCES

- Baars, B. J., Banks, W. P., and Newman, J. N. (eds) (2003). *Essential Sources in the Scientific Study of Consciousness*. Cambridge, MA, MIT Press.
- Banks, W. P. (ed.) (2009). *Encyclopedia of Consciousness*. Academic Press.
- Bayne, T., Cleeremans, A., and Wilken P. (eds) (2009). *The Oxford Companion to Consciousness*. Oxford, Oxford University Press.
- Edelman, D. B., and Seth, A. K. (2009). Animal consciousness: a synthetic approach. *Trends Neurosci.* 32, 476–484.
- Edelman, G. M. (2003). Naturalizing consciousness: a theoretical framework. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5520–5524.
- Koch, C., and Tononi, G. (2008). The neural correlates of consciousness: an update. *Annu. Rev. N. Y. Acad. Sci.* 1124, 239–261.
- Metzinger, T. (2003). *Being No-One*. Cambridge, MA, MIT Press.
- Owen, A. M., Schiff, N. D., and Laureys, S. (eds) (2009). *Coma science: clinical and ethical implications*. *Prog. Brain Res.* 177.
- Seth, A. K. (2009). Explanatory correlates of consciousness: theoretical and computational challenges. *Cogn. Comput.* 1, 50–63.

Received: 12 January 2010; accepted: 25 February 2010; published online: 10 March 2010.

Citation: Seth AK (2010) *The grand challenge of consciousness*. *Front. Psychology* 1:5. doi: 10.3389/fpsyg.2010.00005

This article was submitted to *Frontiers in Consciousness Research* a specialty of *Frontiers in Psychology*.

Copyright © 2010 Seth. This is an open-access article subject to an exclusive license agreement between the authors and the Frontiers Research Foundation, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.