

R E V I E W

---

ELIZABETH IRVINE

*Consciousness as a Scientific Concept: A Philosophy  
of Science Perspective*

Dordrecht: Springer, 2012, 182 pp, £90 (hardback)

ISBN: 978-9400751729

Benjamin Kozuch  
*Philosophy Department,  
University of Arizona  
bigben@email.arizona.edu*

The downfall of structuralist schools of psychology in the early-twentieth century is a well-known piece of cognitive science folklore. The ‘introspectionists’, as their detractors called them, intended to use scientific methods to map the world of sensory experience, but their research programme collapsed somewhat abruptly when it was discovered that the results of the various structuralist labs were incommensurable. If the stronger conclusions of Elizabeth Irvine’s first book, *Consciousness as Scientific Concept*, are correct, we might one day regard contemporary scientific research into consciousness as we presently regard introspectionism: a troubled chapter in the history of cognitive science.

Irvine’s thesis is that the concept of consciousness should be eliminated from scientific practice. This is not only because a science of consciousness cannot bear fruit, but because retention of the concept is an impediment to progress in the cognitive sciences. Irvine uses various lines of attack, arguing that we have no adequate or uncontroversial measure of consciousness (Chapters 2, 3, and 4); that a science of consciousness will not experience the kind of convergence between its measures and methods successful sciences usually do (Chapter 5); that consciousness does not qualify as a scientific kind (Chapter 6); and that a widely used paradigm for research into consciousness, the content-matching method, cannot work (Chapters 7 and 8). The book culminates by bringing together conclusions reached earlier in the book to argue the concept of consciousness meets the conditions under which a concept is usually eliminated from scientific practice (Chapter 9).

This reviewer found the book insightful, timely, and clearly written and organized. It should be a welcome addition to the literature concerning the scientific study of consciousness. While Irvine probably does not succeed in showing ‘there can be no science of consciousness’ (p. 151), she does point out areas in consciousness research that do not yet seem to live up to standards usually at work in science. Furthermore, it is valuable to have, for our perusal, an eliminativist position as well developed as the one in Irvine’s book. The book has many detailed and intricately related parts, and so this review attempts only to provide a schematic of Irvine’s overall argument, and to highlight (and in some places critique) some of the book’s more interesting arguments and ideas.

One of these comes in the first part of the book, where she argues there can be no adequate measure of consciousness, since neither subjective (Chapter 2) nor objective (Chapter 3) measures can work. Putting it roughly, a subjective measure of consciousness counts a subject as conscious of the stimulus if and only if she reports having seen it; an objective measure does so if and only if information about the stimulus is expressed in the subject’s intentional behaviour (Seth *et al.* [2008]). (So, according to an objective measure, success at a forced choice task indicates consciousness, whereas a galvanic skin response does not.) It is not too surprising that a liberal criterion like an objective measure is unworkable, given the counterintuitive results it produces. Notoriously, an objective measure must count the blindsighted subject as conscious of the target even though he insists he cannot see it since he can detect its presence above chance. Irvine has other, novel criticisms of objective measures, but it is her discussion of subjective measures that is more interesting.

Irvine surveys the different subjective indices of consciousness available, from first-order reports, in which the subject is asked to simply report what she thinks she sees (Overgaard *et al.* [2006]; Schwitzgebel [2008]), to more recently introduced measures, such as those that measure consciousness by integrating first-order reports with subjects’ confidence ratings (i.e. how sure the subject is about what she thinks she saw) (Kunimoto *et al.* [2001]). Irvine argues that each of these indices are unusable because of response bias. Response bias occurs when factors unrelated to perception influence a subject’s reports. A subject, for example, might have a bias towards reporting having seen the target because there is a large reward for correctly saying the target appeared, and only a small penalty for incorrectly doing so. Or such a bias might come about because of the training the subject was given. What Irvine argues is that such bias is always present in one form or another, and that this results in relativism about how reports should be gathered and interpreted: ‘Once it is accepted that eradicating bias is impossible [...] many types of bias, and thus many types of report, become viable and appropriate ways of responding to the same stimulus, given an appropriate goal or context’

(p. 28). There is, furthermore, no hope we might ‘train’ away the biases, ‘instead, subjects who have undergone phenomenological training are [merely] equipped with different response biases to the ones they started with’ (p. 27). Since a reliable measure of consciousness seems fundamental to a science of consciousness, the problem Irvine describes seems noteworthy indeed.

Irvine has other insightful critiques of contemporary scientific research into consciousness. She points out how the use of dissociations in consciousness research appears aberrant (Chapter 4). Usually in cognitive science, when dissociations are found, they are in part used to modify the framework used to classify the phenomena being studied. In contrast, the interpretative framework employed in consciousness research rigidly adheres to the conscious/unconscious distinction, regardless of what type of dissociations are found. Irvine also argues that consciousness, as it is sometimes operationalized (Dehaene and Changeux [2004]; Lamme [2006]), does not qualify as a scientific kind (Chapter 6). Take, for instance, global neuronal workspace theory, which holds that content becomes conscious when it is ‘globally broadcast’ to a number of subsystems. Advocates of this theory often operationalize consciousness in terms of reportability, but Irvine argues what is being called ‘reportability’ here is not due to any one mechanism, but rather a collection of mechanisms too disparate to qualify as a scientific kind.

The most contentious part of the book comes in Irvine’s discussion of a popular paradigm for neuroscientific research into consciousness, the content-matching method (Chapters 7 and 8). In the type of content-matching Irvine focuses on, one tries to locate the basis of consciousness by finding some stage of perceptual processing having the same kind of content as is found in consciousness. Block ([2007]), for example, believes consciousness is ‘rich’ in its informational content, and so identifies the contents of consciousness with a large-capacity sensory memory store. Irvine claims this cannot be right, arguing that current models of sensory memory recognize there to be no ‘static, unitary, and detailed’ (p. 124) memory store such as the one Block describes. She dismisses other views of what conscious contents are like on similar grounds (viz. sensorimotor theories (O’Regan and Noë [2001]), and ‘hybrid’ theories positing generic and specific kinds of phenomenology (for example, Kouider *et al.* [2010])).

However, Irvine not only thinks that all attempts at this kind of content-matching have thus far failed, but that all will. She points out how most researchers ‘assume the unity, internal consistency, and non-overlapping nature of conscious content’ (but see Zeki [2003]) and argues this is ‘simply incompatible with many well-known and well-established facts about sensory and cognitive processing, such as the existence of multiple, overlapping and occasionally inconsistent contents of different processing streams’ (p. 146). The idea here seems to be that, if perceptual processing has a fragmented

architecture and is therefore prone to redundant representation, then any plausible characterization of the nature of conscious contents would present them as being disjointed, overlapping, and sometimes inconsistent.

This conclusion is too strong. As far as I can see, the only relevant difference between the contents of the various processing streams and the contents of consciousness (as we typically conceive of them) is that there is more of the former than the latter. This means (not surprisingly) that not all contents in the various processing streams can be conscious. (Otherwise, the contents of consciousness would be disjointed and overlapping.) But this is consistent with a subset of these contents—a subset that is coherent and non-overlapping—matching the contents of consciousness. And so the fragmented, redundant nature of perceptual processing is not itself reason to think no content match is possible, even if a content match is yet to be found.

Moreover, consider a principle widely assumed by researchers of consciousness, which is that if some neural system, *N*, forms the neural basis of experience, *E*, then *N* must have the same content as *E* (cf. Chalmers [2000]). This principle seems but a mundane materialist assumption: it just says any neural system forming the neural basis of, say, an experience as of redness, must itself be representing redness (as opposed to greenness or nothing at all). Now, if such a principle is true, there must be, somewhere among the processing streams, contents matching the contents of consciousness; they will be in whatever processing streams constitute the neural basis of those conscious contents. And so it seems Irvine's claim that there can be no match in content can only be accepted if we have adequate reason to think that the principle under consideration is false. To date, this is something we lack (but see Noë and Thompson [2004]).

The book closes by pulling together earlier parts of the book in service of a general argument for the elimination of consciousness as a scientific concept (Chapter 9). She argues conclusions reached earlier in the book show that the 'methodological norms' usually obtaining in science are not being satisfied in consciousness research and that consciousness meets the conditions (as described in Craver [2007]) under which a concept is usually eliminated from scientific discourse. Taking stock, Irvine decides cognitive science is better off leaving consciousness behind:

An examination of the range of operationalisations, measures, and mechanisms of consciousness show that there is nothing in common between them all, and that instead they refer to a wide range of different phenomena and mechanisms already described in cognitive science [...] it is possible to label these different phenomena as subtypes of consciousness [...] but] having two sets of labels of the same phenomena, one comprised on the vocabulary of perceptual and cognitive abilities, and one with the word 'consciousness' added, is of little practical or explanatory benefit. (p. 154)

What of the idea considered at the beginning of this review? Is it true consciousness research is in such disarray it will one day be regarded—like introspectionism—as another rough patch in the history of cognitive science? This reviewer thinks it is too soon to say. But, if it is, it could well be for reasons Irvine elucidates in her new book.

## References

- Block, N. [2007]: ‘Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience’, *Behavioral and Brain Sciences*, **30**, p. 481.
- Chalmers, D. [2000]: ‘What is a Neural Correlate of Consciousness?’, in T. Metzinger (ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, Boston, MA: MIT Press.
- Craver, C. F. [2007]: *Explaining the Brain*, Oxford: Oxford University Press.
- Dehaene, S. and Changeux, J. P. [2004]: ‘Neural Mechanisms for Access to Consciousness’, in M. Gazzaniga (ed.), *The Cognitive Neurosciences III*, New York: Norton.
- Kouider, S., de Gardelle, V., Sackur, J. and Dupoux, E. [2010]: ‘How Rich is Consciousness? The Partial Awareness Hypothesis’, *Trends in Cognitive Sciences*, **14**, pp. 301–7.
- Kunimoto, C., Miller, J. and Pashler, H. [2001]: ‘Confidence and Accuracy of Near-Threshold Discrimination Responses’, *Consciousness and Cognition*, **10**, pp. 294–340.
- Lamme, V. A. [2006]: ‘Towards a True Neural Stance on Consciousness’, *Trends in Cognitive Sciences*, **10**, pp. 494–501.
- Noë, A. and Thompson, E. [2004]: ‘Are There Neural Correlates of Consciousness?’, *Journal of Consciousness Studies*, **11**, pp. 3–28.
- Overgaard, M., Rote, J., Mouridsen, K. and Ramsøy, T. Z. [2006]: ‘Is Conscious Perception Gradual or Dichotomous? A Comparison of Report Methodologies During a Visual Task’, *Consciousness and Cognition*, **15**, pp. 700–8.
- O’Regan, J. K. and Noë, A. [2001]: ‘A Sensorimotor Account of Vision and Visual Consciousness’, *Behavioral and Brain Sciences*, **24**, pp. 939–72.
- Schwitzgebel, E. [2008]: ‘The Unreliability of Naïve Introspection’, *Philosophical Review*, **117**, pp. 245–73.
- Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M. and Pessoa, L. [2008]: ‘Measuring Consciousness: Relating Behavioural and Neurophysiological Approaches’, *Trends in Cognitive Sciences*, **12**, pp. 314–21.
- Zeki, S. [2003]: ‘The Disunity of Consciousness’, *Trends in Cognitive Sciences*, **7**, pp. 214–8.