

Review of *Understanding Phenomenal Consciousness**

Brad Weslake[†]

29 November 2004

William S. Robinson, *Understanding Phenomenal Consciousness*, Cambridge University Press, Cambridge, May 2004, pp. xii+264, ISBN 0521834635, AUD\$150.00.

THE EPIPHENOMENALISTS RETURN

In recent philosophy of mind, epiphenomenalism—that strain of dualism according to which the mind is caused by the body but does not cause the body in turn—has undergone something of a renaissance. Contemporary epiphenomenalists bear only partial resemblance to their more extravagantly metaphysical ancestors, however. Traditional epiphenomenalists thought that (at least) two sorts of mental properties were epiphenomenal—intentional properties such as the meaning or representational content of the propositional attitudes (beliefs, desires and so on); and conscious properties such as awareness and the qualitative nature of experience. Contemporary epiphenomenalists, on the other hand, are largely sanguine about the prospects for *intentionality* to be brought within the purview of a physicalist worldview; what forces their dualism is one particular feature of *consciousness*—what irks them are *qualia*, the

*Published in *Metapsychology Online Reviews*, Vol. 8, No. 49, 29 November 2004. Online: <http://www.mhnet.org/books/books.php?type=de&id=2411>. Thanks to Christian Perring for supplying the review copy, and to Nandi Theunissen for discussion.

[†]Centre for Time
Department of Philosophy
Main Quad A14
University of Sydney
NSW 2006
Australia
brad.weslake@arts.usyd.edu.au
<http://www.usyd.edu.au/time/weslake/>

“what it is like” of experience (Nagel, 1974), or in the idiom adopted by Robinson for the title of his book, what philosophers refer to as *phenomenal consciousness*.

William Robinson’s *Understanding Phenomenal Consciousness* is a tightly argued and original defence of a form of epiphenomenalism about consciousness he terms *Qualitative Event Realism* (hereafter *QER*). The most famous contemporary arguments for epiphenomenalism, the knowledge argument due to Frank Jackson (see Ludlow, Nagasawa, and Stoljar, 2004) and the conceivability argument due to David Chalmers (1996), have spawned a voluminous secondary literature, and it is a virtue of Robinson’s book that he leaves these to one side, in order to form his own line of attack. Robinson is, like other epiphenomenalists, concerned primarily with qualia; but perhaps uniquely among them, he thinks that the route to epiphenomenalism lies not via knockdown *a priori* arguments concerning what can in principle be integrated within a physicalist worldview, but rather via careful evaluation of physicalist theories of consciousness, in light of the current state of neuroscience. Part I of the book forms the philosophical core, where Robinson criticises rival views and motivates his own; while Part II consists of a slightly more speculative sketch of how a science of consciousness, coupled with the metaphysical stance of epiphenomenalism, might proceed (that such a science is possible might seem a surprising claim, but Robinson’s version of epiphenomenalism is, he argues, empirically indistinguishable from a sufficiently fine-grained functionalism—the metaphysical difference being whether the mind-body relation is one of causation, or identity).

The book overall is aimed primarily at a philosophical audience; it is densely populated with argument, and is at times reasonably technical. I expect it would be challenging for those unfamiliar with the landscape of recent philosophy of mind, though Robinson writes with such clarity and precision that it could serve as a useful (though idiosyncratic; see below) introduction to some of the territory in the philosophy of consciousness, for those already in the neighbourhood. In particular, I found his arguments for a range of theses concerning the subjective properties of phenomenal experience to be exceptionally lucid (notwithstanding their alleged literal spatiality). These theses are for the most part independent of the argument for epiphenomenalism. Also especially good are his criticisms of representationalist and higher order thought (HOT) theories of consciousness in Chapters 4 through 7—readers who are ontologically impressed by qualia in the first place and yet subscribe to one or another of these physicalist theories will find their allegiance strongly questioned. But while Robinson is largely convincing on the negative side of his metaphysical project, he is much less so on the positive side. In the remainder of this review I will restrict myself to two themes—firstly the architecture of the book, and secondly a brief criticism of Robinson’s solution to the causal problem for dualists.

EXPLAINING QUALIA

As I have said, Robinson (profitably, I think) eschews engaging with some of the most popular arguments in the recent literature on consciousness. However the framework in which he sets the discussion is somewhat idiosyncratic. Right from the beginning of the book, the axis around which the argument turns is the debate between “experiential realists” (those who believe in qualia) and “minimalists” (those who do not). But as Robinson himself (p. 10) recognises, very few philosophers are minimalists (or at least, very few who work on these problems). The primary debates here are over the metaphysical *status* of qualitative events, not their *existence*. Setting up the debate in this way has the happy result that his QER, though a form of epiphenomenalism, turns out to be just one variant of experiential realism—which, in turn, makes it seem as if motivating it is a matter of merely working out some details; of merely resolving some in-house disputes with others in his camp—when in fact, most physicalists and (presumably) all dualists are camped right there with him. Which is to say that setting up the debate in these terms serves to rhetorically deflate what are generally taken to be the most important debates in the area.

I take it this is intentional. Firstly, it is the sheerly phenomenal character of qualia (independent of their status) that form the core of his criticisms of his physicalist rivals. Secondly, Robinson does not buy into the various metaphysical debates over physicalism and dualism. Neither physicalism nor dualism are defined, and supervenience—the metaphysical relation around which much mind-body debate has revolved in the last forty or so years—is left off the conceptual map altogether in the early chapters, only making it onto the stage by Chapter 8, or more than halfway through the book. It is only at this point in the book, too, that we get an explanation of why this is so. Robinson argues that just as posing a theoretical identity in absence of an explanatory relation is empty, so is posing supervenience in absence of an explanation of the necessity involved. This is a common line of argument, and vindicates somewhat the way the debate is set up. But of course the epiphenomenalist has no good explanation, of the sort demanded, either—and therefore placing the emphasis on explanation would seem to count against *any* current theory of consciousness. What the debate becomes, once this is clear, is a matter of what we ought to say were we to have some form of explanation (something of the sort Robinson outlines in Part II, say)—and on these terms it is, at the least, not clear that epiphenomenalism is left with any advantage. Indeed, I think it is left at a disadvantage, for reasons which follow.

KNOWING YOUR OWN MIND

In Chapter 10, Robinson arrives at the hard question for the epiphenomenalism of his QER: How could it be that we have knowledge of our phenomenal qualities if they are inefficacious? The core objection here can be simply stated as follows:

Consider two worlds, one where epiphenomenalism is true; and another where there are no qualia at all, but that is otherwise a physical duplicate of the first. How could a person in either world know which world they are in?

The answer Robinson gives to this question is all-too-brief, and in any case unsatisfying (in addition, coming late as it does, it turns out to be narratively disappointing, given the strength of the earlier sections of the book). He notes that *if epiphenomenalism is true*, our believing and speaking about phenomenal events counterfactually depends on the phenomenal events themselves, since the neural causes of our relevant beliefs and language are also the causes of the phenomenal events. For example, we wouldn't ever (truly) report "I see a purple haze" unless we did see a purple haze, since one and the same neural event is the cause of both the purple haze and our report. And that is the whole account—remarkably, Robinson claims that this is all that is needed to ground our knowledge of phenomenal properties. But of course, counterfactual dependence of this sort is radically insufficient for knowledge. On this model, for example, we should be able to do our neuroscience from the armchair—since our beliefs and language presumably have precisely the same counterfactual dependence relation with their neural causes as they do with the phenomenal properties themselves (indeed, Robinson's account of knowledge would be perfect for the identity theorist, for whom purple hazes *just are* neural events).

As Robinson (p. 169) recognises, it might also be argued that his argument here is question begging, in virtue of the conditional ("*if epiphenomenalism is true...*") as highlighted above—surely you cannot argue for epiphenomenalism with a premise conditionalising on its truth? The reply Robinson gives to this objection itself turns on the account of knowledge just criticised. What Robinson says, in effect, is that we are entitled to believe in epiphenomenalism as an argument to the best explanation, as follows:

1. We have knowledge of phenomenal events.
2. Epiphenomenalism is the best account of (our knowledge of) phenomenal events; therefore
3. Epiphenomenalism is true.

But again, the account of our knowledge of epiphenomenal events given by Robinson is not up to the task. The problem, then, is simply that epiphenomenalism does not

in fact provide the best account of phenomenal consciousness, since it leaves it mysterious how we could ever have knowledge, form memories, and otherwise cognitively access phenomenal properties. Midway through the chapter, Robinson includes a meditation worthy of Descartes, where he grapples with himself on this point—with a view to convince himself (and us) that he is justified in believing in phenomenal events by the lights of his view, and therefore that his account is defensible. And he is in a real bind. On the one hand, the book is a sustained defence of epiphenomenalism; on the other, he plainly knows he is undergoing phenomenal experiences. But how? The conclusion seems to be that from these facts alone, there is no contradiction in supposing epiphenomenalism compatible with our possessing knowledge of phenomenal events. But whatever *other* attractions epiphenomenalism may have, they can't help *this* part of the story—the gap needs filling.

Earlier in the book, as I have said, much is made of the fact that physicalist accounts of phenomenal consciousness seem unable to *explain* qualia; once Robinson has given his own account, the natural question is whether he has improved the situation. Does establishing a causal connection between certain neural events and certain qualitative events really *explain* qualia? Does preserving their separate metaphysical status somehow provide a stronger explanation than reducing them to the very same neural events? It is hard to see that this is a difference in explanatory power, rather than a matter of whether one's metaphysics is seen to reflect the perceived specialness of the target domain. Indeed, once we see that explanatory force is equally a problem for the epiphenomenalist, we might think the physicalist is actually better off—after all, she has a story about how the legendary explanatory gap could be expressive not of an ontological quandary, but simply of the fact that we don't yet have any explanations.

REFERENCES

- Chalmers, David John. 1996. *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press, New York, 1996. [2](#)
- Ludlow, Peter; Nagasawa, Yujin; and Stoljar, Daniel (Eds.). 2004. *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, MIT Press, Cambridge MA, 2004. [2](#)
- Nagel, Thomas. 1974. "What Is It Like to Be a Bat?", in *The Philosophical Review*, Vol. 83(4), October 1974, pp. 435–450. [2](#)