

Derk Pereboom

Consciousness and the Prospects of Physicalism.

Oxford and New York: Oxford University Press 2011.

208 pages

US\$65.00 (cloth ISBN 978-0-19-976403-7)

Is physicalism true? And if so, can it be of a nonreductive variety? Consciousness presents an obstacle to an affirmative answer to the first question. The phenomenal properties of consciousness do not seem physical, and this appearance has motivated sophisticated arguments against physicalism. Regarding the second question, the challenge is to say how the phenomenal features of consciousness, as well as other mental phenomena, can be intimately enough related to the physical domain for physicalism to be true, and yet irreducible to purely physical phenomena in a philosophically significant way. Derk Pereboom's *Consciousness and the Prospects of Physicalism* is an engaging, carefully argued, attempt to address these issues.

Chapters 1 through 6 are devoted to showing how physicalism might be true and, in particular, how a physicalist should address the well-known 'knowledge argument' of Frank Jackson and the 'conceivability argument' recently defended in great detail by David Chalmers. Pereboom suggests two strategies worth pursuing. The first, developed at length in Chapters 1 through 4, is the 'qualitative inaccuracy hypothesis' according to which 'introspection represents phenomenal properties as having a certain qualitative nature, which these properties may lack' (3). One way of spelling this out—and one that Pereboom takes to merit serious consideration—is the 'primitivist inaccuracy hypothesis' that we 'either represent phenomenal properties as primitive or else our introspective representations of phenomenal properties are primitive-belief occasioning' (i.e., given how we introspectively represent phenomenal properties, we have a strong tendency to believe that they are primitive) (16-17). Perhaps, however, phenomenal properties are not primitive, and in this case our introspective representations of phenomenal properties will be inaccurate. If, as Pereboom contends, the knowledge argument and the conceivability argument presuppose that our introspective representations of consciousness are accurate, then if it is an open possibility that such representations are inaccurate, it will follow that such arguments fail to establish the falsity of physicalism. The more general thought is that it is possible to draw an appearance/reality distinction for the phenomenal properties of consciousness. If such a distinction can be drawn, the fact that phenomenal properties do not appear to be physical will not mandate the conclusion that they are not physical. This thought has been advanced under somewhat different guises elsewhere in recent literature (see, for example, Christopher Hill's *Consciousness*), and must be addressed by advocates of both the knowledge argument and the conceivability argument.

In advancing the qualitative inaccuracy hypothesis, Pereboom develops a useful analogy with secondary properties, especially color. It is plausible, he suggests, that our sensory experience represents colors as primitive, unanalyzable properties of objects; yet it is also plausible that there are 'no instantiated primitive color properties' (36). Thus, there may be an important sense in which sensory experience misrepresents color.

However, Pereboom argues that with both color properties and phenomenal properties we can accept inaccuracy theses without having to endorse eliminativist conclusions. He bases this on the view that ‘the structure of certain concepts is a conjunction of conditionals’ (31), and deems it an open possibility that, e.g., analysis of ‘phenomenal redness’ will reveal a conjunction of conditionals, roughly as follows (31-40):

P1. If there is a property that resembles the introspective representation of phenomenal redness, then the concept ‘phenomenal redness’ applies to that property.

P2. If there is not a property that resembles the introspective representation of phenomenal redness, but there is a property that normally causes such introspective representations, then the concept ‘phenomenal redness’ applies to that property.

P3. If there is neither a property that resembles the introspective representation of phenomenal redness, nor a property that normally causes such introspective representations, then the concept ‘phenomenal redness’ applies to whatever properties cause or could cause the introspective representation of redness.

If the structure of phenomenal concepts is something like this, then the failure of phenomenal properties to resemble how they appear in introspection will not entail that phenomenal concepts fail to correctly apply to anything. Somewhat differently, if the qualitative inaccuracy hypothesis is true, this will support eliminativism about phenomenal properties only if we suppose that the concept of phenomenal redness, say, can be correctly applied only to a property that resembles the introspective representation of phenomenal redness. Thus, an argument from qualitative inaccuracy to eliminativism about the phenomenal will have to include an argument against phenomenal concepts having the richer kind of structure sketched above. As there may be colors even if we misrepresent them, there may be phenomenal properties even if they are not as they seem in introspection.

Should the qualitative inaccuracy hypothesis fail, Pereboom contends that a physicalist may instead hold that we are ignorant of the categorical or intrinsic nature of physical properties, and that this nature of which we are ignorant is what is needed to provide a physicalist or materialist account of consciousness. This outlook—‘Russellian Monism’—is developed in Chapters 5 and 6, where Pereboom argues that there may indeed be such properties, and that they may be physical properties, though ‘protophenomenal’ in that they account for the phenomenal features of consciousness. As Pereboom seems to recognize, and as others have noted, such positions essentially appeal to properties that are at least presently unconceived, and we may worry that an appeal to protophenomenal properties thus amounts to little more than the affirmation that there is something non-phenomenal in virtue of which phenomenal properties are instantiated. Nonetheless, if physicalist Russellian Monism is at least coherent, a successful argument against physicalism will have to rule against the possibility that while physical properties are responsible for phenomenal consciousness, we are ignorant of these properties, and this is why we lack a satisfactory physicalist account of the phenomenal.

Chapters 7 and 8 defend what Pereboom calls ‘robust nonreductive physicalism’. The first component of this defense consists in rejecting all identifications between the microphysical on one hand and the mental on the other. While others have rejected ‘token identity’—the identity of each instance of a mental property with an instance of a physical property—Pereboom goes further, arguing that there are causal powers that are not microphysical. Here Pereboom targets an alternative nonreductive outlook according to which while instances of mental properties are not identical to instances of microphysical properties, it is nonetheless the case that every causal power of a mental property is a causal power of its microphysical realizer (see, for example, Sydney Shoemaker’s *Physical Realization*). Concerning such an outlook, Pereboom argues that even at the level of token causal powers, the closest relationship between the mental and the microphysical is constitution. And yet this, he contends, is enough for physicalism.

The second component of Pereboom’s robust nonreductive physicalism, defended in Chapter 8, consists in the rejection of classic functionalism as an adequate basis for a nonreductive outlook. Here Pereboom agrees with critics of nonreductive physicalism, Jaegwon Kim in particular, that it is not easy to see how instances of functional properties—properties the having of which consists in having some other property that plays a certain role—could have causal powers that are anything but purely physical causal powers. Instead, he suggests that the nonreductive physicalist should hold that mental properties, including phenomenal properties, are ‘compositional properties’, ‘properties things have solely by virtue of intrinsic features of their parts, either proper or improper, and relations these parts have to one another’ (148). By taking mental properties to be compositional properties in this sense, Pereboom hopes to secure the nonreductive desideratum that ‘the mental...be causally efficacious qua mental’ and not ‘just by virtue of being realized by microphysical property instances that are causally efficacious’ (148). In other words, we can say how minds may have novel causal powers, yet do so in a way compatible with physicalism.

Pereboom’s defense of nonreductive physicalism is well-constructed and comprehensive. While there are several points that a reductionist may take issue with—including the apparent identification of reductionism with microreductionism, the specific arguments against token identity, and the relationship between compositional properties and microphysical properties—Pereboom succeeds in developing a novel outlook, one that cannot be faulted on the same grounds as traditional nonreductive frameworks.

Consciousness and the Prospects of Physicalism is an ambitious, often subtle, approach to the contemporary debate over the status of physicalism. It merits careful attention, and will have to be grappled with both by those who contend that physicalism is false as well as those who think that any physicalism has to be of a reductive variety. It advances the debate in a compelling and original way, and will be of interest to anyone working on the issues that Pereboom discusses.

Kevin Morris
Tulane University