

More Troubles for Epiphenomenalism

Hans Muller

Received: 6 August 2008 / Accepted: 11 August 2008 /
Published online: 11 November 2008
© Springer Science + Business Media B.V. 2008

Abstract I have argued that to say qualia are epiphenomenal is to say a world without qualia would be physically identical to a world with qualia. Dan Cavedon-Taylor has offered an alternative interpretation of the commitments of qualia epiphenomenalism according to which qualia cause beliefs and those beliefs can and do cause changes to the physical world. I argue that neither of these options works for the qualia epiphenomenalist and thus that theory faces far more serious difficulties than has previously been recognized.

Keywords Qualia · Epiphenomenalism · Causation · Frank Jackson

The debate between Dan Cavedon-Taylor and me reveals that the qualia epiphenomenalist faces a hard choice: either qualia cannot cause beliefs or qualia are not epiphenomenal. At one level, our dispute is over what commitments follow from accepting the proposition that *qualia are causally impotent with respect to the physical world*. But at another, deeper, level this debate is over the question of whether there is any coherent and consistent way to describe a subjective, felt realm of mental properties that are caused by, but do not cause, physical processes. For reasons I explain below, I do not think Cavedon-Taylor's defense of Frank Jackson's version of qualia epiphenomenalism is successful. However, I think Cavedon-Taylor's reinterpretation of the commitments of that view is worth careful consideration because it brings into sharp focus how exceedingly narrow a space is left for mental properties when one extends the scope of "physical" to "a wide sense of 'physical' which includes everything in *completed* physics, chemistry, and

H. Muller (✉)

Department of Philosophy, American University of Beirut, Beirut, Lebanon
e-mail: hm38@aub.edu.lb

neurophysiology, and all there is to know about the causal and relational facts consequent upon all this, including of course functional roles.”¹

Since there are many, many causal and relational facts about beliefs and since they have functional roles, it seems clear that beliefs must count as physical according to the qualia epiphenomenalist. But this means that beliefs cannot be caused by qualia on pain of violating the following restriction: “the instantiation of *qualia* makes a difference to *other mental states* though not to anything physical.”²

None of these data about the relationship between qualia and beliefs were apparent to me before reading and responding to Cavedon-Taylor’s objection to my argument. So I recommend the details of that exchange to readers interested in the prospects of qualia epiphenomenalism. Here is the core of Cavedon-Taylor’s defense of Jackson against my argument:

It seems to me that Jackson can respond to Muller’s argument by claiming that his qualia caused him to have the belief that he possesses qualia and that it is this later, *belief*, state that is the direct cause of his writing the paper in question. On this explanation, the epiphenomenal status of qualia is preserved insofar as such states fail to directly causally interact with the physical world; rather, Jackson’s *belief* in the existence of qualia directly causally interacts with the physical world by bringing about, perhaps with the help of other attitudinal states, an assertion of qualia’s existence.³

So the picture is that a quale leads to the belief that the quale in question exists and the belief leads to the assertion that qualia exist. This move involves a rejection of the criteria I implicitly relied on in my argument: i.e., the thesis that to say something is causally inefficacious with respect to the physical world is to say that if that thing ceased to exist, then the physical world would remain unchanged.

I note that in my original article and in Cavedon-Taylor’s reply to it, the discussion was focused largely on the things Jackson wrote while trying to allay fears about his position under the heading of “the bogey of epiphenomenalism.” But it would be a mistake to forget that in addition to responding to objections, Jackson also gave an argument. Namely, the knowledge argument, which appeals jointly to the existence and the physical causal impotence of qualia to argue that physicalism is false.⁴ This is the set-up of the thought-experiment on which that argument turns:

Mary is confined to a black-and-white room, is educated through black-and-white books and through lectures relayed on black-and-white television. In this way she learns everything there is to know about the physical nature of the world. She knows all the physical facts about us and our environment, in a wide sense of ‘physical’ which includes everything in *completed* physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this, including of course functional roles. If physicalism is true, then she knows all there is to know. For to suppose

¹ Jackson (1986: 291). Emphasis in the original.

² Jackson (1982: 133). Emphasis in the original.

³ Cavedon-Taylor (2008: 106). Emphasis in original.

⁴ I am indebted to Joe Neisser for forcefully emphasizing this point in discussion.

otherwise is to suppose that there is more to know than every physical fact, and that is just what physicalism denies.⁵

When she is freed from her surreal philosophy-example captivity and steps out of the room, she sees colors for the first time. And the experience of seeing color, involving as it does a quale, provides genuine knowledge of a non-physical fact and thus falsifies physicalism. Jackson makes the following clarification about just what is at issue in his argument:

The whole thrust of the knowledge argument is that Mary (before her release) does *not* know everything there is to know about brain states and their properties, because she does not know about certain qualia associated with them.⁶

But if the qualia epiphenomenalist takes on Cavedon-Taylor's friendly amendment, it seems that pre-release Mary *would* know about the qualia that are associated with particular brain states. To see why this is so, consider the fact that the scope of Mary's knowledge could extend to the states of Frank Jackson's brain in 1982. If she were asked, having mastered "*completed* physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this, including of course functional roles," whether the owner of that brain would produce an article defending the existence of qualia and attacking physicalism she would say "yes". Why? Because she would be able to produce a functional map of the brain states which would produce such behavior. And some of those brain states, namely the ones instantiating the belief-in-qualia, would only exist if Jackson in fact had qualia. In this way we can see that the adoption of Cavedon-Taylor's amendment would undermine the knowledge argument for physicalism.

So when Cavedon-Taylor writes, "All Jackson, *qua* qualia epiphenomenalist, should be read as committing to is the claim that qualia do not *directly* cause anything in the physical world. This commitment is compatible with many things,"⁷ he is certainly correct. One of the many things it is compatible with appears to be the version of physicalism that Jackson was trying to undermine in "Epiphenomenal Qualia." And that, I submit, is a very odd place for a qualia epiphenomenalist to end up.

That last point sums up the reasons I think Cavedon-Taylor's friendly amendment to qualia epiphenomenalism violates the spirit, if not the letter, of that theory. However, his proposal does create an opening for a position which fits qualia – *qua* subjective, felt properties of brain states – into a physicalist ontology. Jackson's original proposal was to make each quale a sort of causal dead-end. That is, as properties that supervene on various brain states they are caused by physical processes but are explicitly said not to be capable of causing any physical processes (including brain processes). But according to the proposal that began with my critique of Jackson and that has been so ably developed by Cavedon-Taylor, qualia are now seen as capable of causing beliefs which – *qua* functionally specifiable

⁵ Jackson (1986: 291).

⁶ Jackson (1986: 291).

⁷ Cavedon-Taylor (2008: 107).

states governed by the laws of a hypothetical complete cognitive neuroscience – are solidly in the realm of what Jackson would consider *physical*.

But this brings us up against one of the really difficult peculiarities of Jackson's qualia epiphenomenalism. To what, exactly, is one committing when one says, "it may be that you have to hold that the instantiation of *qualia* makes a difference to *other mental states* though not to anything physical."⁸ What is the ontological status of these mental states? We are told they are non-physical, but for Jackson the scope of "physical" covers all entities with specifiable functional roles.⁹ So now the qualia epiphenomenalist has to countenance mental states that are non-physical and have no functional role. But what states are these? According to Jackson qualia are non-physical and have no functional role, but they aren't mental *states*. Recall that in his effort to differentiate his qualia epiphenomenalism from what he called "the classical epiphenomenalist position" he explicitly rejected the thesis that "mental *states* are inefficacious with respect to the physical world" in favor of the following view:

All I will be concerned to defend is that it is possible to hold that certain *properties* of certain mental states, namely those I've called qualia, are such that their possession or absence makes no difference to the physical world.¹⁰

Since qualia are properties as opposed to states, they cannot be the non-physical states with no functional role that the theory asks for. And it is not at all clear what sort of state can answer to this job description. And *that* is the truly problematic feature of qualia epiphenomenalism.

So it seems that the advocate of Jackson's version of qualia epiphenomenalism is faced with a nasty dilemma. One option is to stand firm on the thesis that qualia have neither direct nor indirect causal power with respect to the physical realm. But as he have seen, this would mean that asserting the existence of qualia falsifies the theory. The second option is to accept Cavedon-Taylor's amendment to the theory and accept the idea that qualia can cause mental states that in turn cause behavior and other physical changes. But as we have seen, this runs afoul of Jackson's claim that things with specifiable functional roles are physical things. Cavedon-Taylor's proposal is useful for expanding our understanding of the possible variations on qualia epiphenomenalism, but in the end the prospects for saving the theory from my original objection seem faint indeed.

References

- Cavedon-Taylor, D. (2008). Still Epiphenominal Qualia: Response to Muller. *Philosophia*, 37(1), 105–107. doi:10.1007/S11406-008-9152-6.
- Jackson, F. (1982). Epiphenomenal Qualia. *The Philosophical Quarterly*, 32, 127–136. doi:10.2307/2960077.
- Jackson, F. (1986). What Mary didn't know. *The Journal of Philosophy*, 83, 291–295. doi:10.2307/2026143.
- Muller, H. (2008). Why qualia are not epiphenomenal. *Ratio*, 21, 85–90.

⁸ Jackson (1982: 133). Emphasis in original. This claim of Jackson's is also cited by Cavedon-Taylor as support for his interpretation of the commitments of qualia epiphenomenalism.

⁹ Jackson (1986: 291).

¹⁰ Jackson (1982: 133).