

NOT SO EPIPHENOMENAL QUALIA

OR

HOW MUCH OF A MYSTERY IS THE MIND?

Fredrik Stjernberg

Växjö University

In recent philosophy of mind, physicalism - in one form or another - has been the dominant position. And this is no surprise. The obvious alternative, dualism, is still struggling to come to terms with a problem that was already presented to Descartes: If the material world can be accounted for by physics, how could something mental, i.e., non-material, come to influence the material world? In the *Objections V*, Gassendi expressed this worry: “How can there be effort directed against anything, or motion set up in it, unless there is mutual contact between what moves and what is moved? And how can there be contact without a body?”

The general idea behind this worry has also been known as the claim of the *causal closure* of the physical. The physical world is causally closed; nothing outside physics can influence

events in the physical world. Given the nature of the resulting problem, it is not unexpected that Descartes' replies were somewhat disappointing. And this problem is still with us.¹

But physicalism is not without its own problems. It seems that the hardest problem for physicalism is consciousness, the subjective aspect of things, the way things seem to a conscious subject.² This is something which we all know a lot about from our own experience, but as yet we don't really have any ideas about how it is to be dealt with scientifically or even systematically. As a prominent theorist in the philosophy of mind, Ned Block, put it recently:

“[I]n the case of consciousness, we have nothing - zilch - worthy of being called a research programme, nor are there any substantive proposals about how to go about starting one. ... Researchers are *stumped*. There have been many tantalizing discoveries recently about neurophysiological syndromes in which consciousness seems to be in some way missing or defective ..., but no one has yet come up with a theoretical perspective that uses these data to narrow the explanatory gap, even a little bit.”
(Block (1994), p. 211)

If we cannot incorporate facts about consciousness into a physicalist picture of the world, then physicalism will be leaving something out, and physicalism will turn out to be false. (No one denies that *some* events can be accounted for by physics, but physicalism is inherently imperialist: if physics is in principle unable to account for some phenomena, then physicalism is false, period.) So, are there facts which physicalism misses?

Frank Jackson's thought experiment with Mary, the colour-deprived colour scientist, is intended to make the case that there are facts which physicalism cannot incorporate, namely facts about *qualia*.³ Jackson's thought experiment is somewhat indirect, in that it is designed

¹For the extensive recent debate, the reader can turn to Heil and Mele (1993), Yablo (1992).

²Other important issues for physicalism are accounting for intentionality and normativity, along with the question whether there is an acceptable and non-trivial formulation of physicalism. These issues can be left for some other occasion.

³See Jackson (1982) and (1986).

to show that there is knowledge which goes beyond all possible physical knowledge, hence that physics cannot tell us everything there is to know, hence that physicalism is false.

The thought experiment is as follows. Mary is a brilliant physicist, and finds out everything that physics could tell us about colours; she constructs a complete physical account of colour and colour perception. She knows everything there is to know about colour from a physicist's point of view. This point of view does not only encompass current knowledge about colour: we can fancy an increase in our physical knowledge that is as large as we wish, as long as it is legitimately construed as physical knowledge. Mary has one disadvantage compared with normal contemporary subjects, however: she has never had any contact with coloured objects. She was raised in a colour-deprived environment, so all she knows about colour and colour-perception is arrived at without any first-hand acquaintance with coloured objects and images. One day she is let out of her black and white, colour-deprived, room, and enters into the normal coloured world. The question is then whether Mary, when first getting out of her room and seeing coloured objects, learns some new facts, acquires some new knowledge, knowledge about what it is like to see coloured objects. Now she for instance knows first-hand what it is like to see a red object; she has the appropriate qualia.

Jackson's contention is that Mary, when she steps out of the colour-isolated room, learns something new, something about how colours *look*. Mary knows all there is to know about physics. But by having the relevant qualia, she learns something new, something that no amount of physical theory could have taught her. Therefore, physicalism is false, since complete knowledge of what physicalism states does not amount to complete knowledge of all the facts. Jackson goes on to argue that qualia are epiphenomenal, and he rebuts several objections to that conception of qualia. If qualia are epiphenomenal, then the ideal of causal closure for the material is maintained, but physicalism is incomplete; it misses some facts.

Physicalists have attempted to counter this understanding of the thought experiment in different ways. Some have claimed that what Mary acquires when stepping out of her black

and white room is not factual knowledge, but an ability, such as the ability to apply a certain concept.⁴ It need not be part of physicalism to hold that any amount of physical information automatically gives the subject a certain ability. Full knowledge of obstetrics won't make you pregnant.

Yet other physicalists hold that there may be new factual knowledge in the situation described, but that the facts known could in principle be known in some other - physicalistically acceptable - way.⁵ According to this strategy, the shortcoming of physicalism lies only in physicalism's failure to give an account of all possible modes of presentations of all facts.

A third physicalistic strategy is Dennett's: cast doubts on the very notion of qualia.⁶ Derk Pereboom suggests a fourth reply in a recent paper.⁷ He holds that, even if physicalism is accepted, there will be no reason to assume that mental states actually are as they appear to introspection. Physicalism will only be understood as asserting that we could have knowledge about "the real nature of a subject's mental states", but not necessarily about "a subject's mental states, as they are apprehended in introspection".⁸

All of the above physicalistic strategies work, to some extent, but they also encounter problems. The first strategy, which leans on a strict division between factual knowledge and abilities, may have some difficulties in spelling out a clear-cut distinction between the two, and it seems clear that Mary *does* acquire some new factual knowledge, not just a new ability. The second strategy is open to the charge that knowledge about these different modes of

⁴Nemirow (1990), Lewis (1990).

⁵Churchland (1985), Bigelow and Pargetter (1990).

⁶Dennett, (1990).

⁷Pereboom (1994).

⁸Pereboom, p. 328.

presentation is in itself an example of knowledge that transcends physicalistically expressible knowledge, so that the physicalist has simply not dealt with the issue properly. The third strategy is difficult to assess, and its relevance can also be questioned. It would take further argument to show that physicalism is vindicated, if there are no qualia. The thought experiment may point to some phenomenon that still cannot be accounted for in physicalistic terms, even if there are no qualia - the physicalists may simply be attacking an unnecessarily strong conception of qualia. Pereboom's suggestion merits further discussion elsewhere, but it still points to a certain limitation in physicalism; the subjective aspect of things, how things seem to the subject, is simply left out of the account.

In this paper, I will be discussing a further physicalist response to the thought experiment. This physicalist response has not been discussed in the literature, and it is perhaps simpler than the other ones discussed in the literature (some might want to call it simplistic).⁹ At the same time, this response leads to further difficulties, and it seems that we in fact face a real dilemma in the philosophy of mind, and some ramifications of this dilemma will be discussed briefly.

The simplistic argument against Jackson's view is as follows. If qualia are epiphenomenal, then Mary could not have had knowledge about them, and she could not have perceived them; knowledge and perception are notions interwoven with notions of causality, and epiphenomena (per definition) lack causal powers. Therefore, if she acquires new knowledge when leaving her secluded room, what she has knowledge of must be endowed with at least some causal powers, or else she has perceived nothing new, learned nothing new, hence qualia cannot be epiphenomenal. And, continues the physicalist, if she acquires new knowledge, it is to be accounted for by physicalism, since it has causal powers, and is therefore part of the physical order of things.¹⁰

⁹Braddon-Mitchell & Jackson come close to discussing it, (1996), p. 134.

¹⁰There is an obvious gap in this reasoning, from "having causal powers" to being "part of the physical order of things". The gap is discussed below, but in general, the idea that something could have clearly causal powers on some non-physical part of nature (a guiding idea of dualist theories) is not very attractive.

Therefore, physicalism stands, at least against this attack, or else the idea that qualia are epiphenomenal has to be given up. This objection may be of some interest to the hardline physicalist, since it retains more of the physicalist position than many other ways of dealing with Jackson's thought experiment. The intention is not to provide a knock-down objection againsts Jackson's interpretation of his thought experiment. The intention is rather to indicate another way in which the physicalist could react to the thought experiment, and in this paper I will spell out some of the ramifications of this reaction. There are several ways to meet the objection, and I will discuss them in turn.

Even if the physicalist stands a better chance of handling Jackson's thought experiment in the way suggested, a problem remains. This is that if qualia are not epiphenomenal, it is difficult to see what status they might have - and our very strong intuitions in favour of Jackson's preferred construal of the thought experiment are still with us. If qualia are needed for a full account of the thought experiment, and they cannot be epiphenomenal, then it remains a mystery how the situation is to be accounted for. The physicalist is reduced to saying "there must be an account of this situation". Is this reaction just dogmatism? Not necessarily: here we have a combination of an argument for the general existence of an account, plus a general inability to understand how such an account is to be constructed.¹¹

One thing, at least, should be answered in the physicalist's challenge. This is that without an adequate account of how a subject is to be granted knowledge of things like qualia, a bare appeal to knowledge of qualia does not accomplish anything; without the extra account, we simply cannot take it for granted that the subject actually *has* knowledge about the supposed entities, qualia.¹²

¹¹The situation is thus similar to that described in McGinn (1991).

¹²See section VII below, for more on this.

Look again at the thought experiment. The whole description of what happens when Mary leaves the laboratory, and her resulting change in knowledge, is steeped in causal terms. Mary comes to realize that red tomatoes look like *this*, she *learns* that this and this is the case. Her acquaintance with red objects gives rise to new experiences, new qualia, and as a result of these qualia, she knows something she didn't know before. Holding that the qualia are the results of new events, hence caused, is consistent with holding that they are epiphenomenal,¹³ but the next step, that these qualia in their turn cause anything, is not, and the way in which the thought experiment is set up rests on endowing qualia with causal properties.

I

The first reaction would be to question the account of knowledge that lies behind the objection. Why should there be this strong connection between knowledge and causation, and why should epiphenomenal qualia pose any special difficulties? Since it is perfectly obvious that we are acquainted with qualia - at least according to the "qualia freak"¹⁴ - this obvious *datum* is something our account of knowledge should take into consideration, and not something we could reject on the strength of some controversial account of the nature of knowledge.

The general objection against a connection between knowledge and causality may very well have some force, but this is not really the right place to assess it in any greater detail. The point of the original response to Jackson's thought experiment was simply to spell out what some physicalists - who certainly should be expected to accept some kind of causal view of knowledge - should say about Jackson's thought experiment; I am not trying to provide a full account of what physicalism should amount to, and then defend that account.

¹³Actually even this step could be questioned: if causality is, in one way or another, tied with the transfer of energy, it would probably be another violation of physics to have things that are caused, but cannot cause anything else.

¹⁴This term, being Jackson's own, is not intended to be abusive.

If we leave aside questions about the general justification of a causal conception of knowledge, there are still two kinds of case that should be discussed. These two are cases where there is no causal tie, but most people want to accept that we have knowledge. These cases, which might indicate limitations to a causal conception of knowledge, might be used - perhaps our suggested understanding of qualia could be understood on these lines? The first proposed exception is that we can have knowledge about things that don't actually have any causal impact on us, we can perhaps know things about the future, or about things outside the light cone (Braddon-Mitchell & Jackson, p. 134). The second exception is abstract objects. It is at least *prima facie* likely that we know some mathematical truths, but abstract objects don't have any causal powers at all. This case will be discussed in the next section.

As for the first case, this couldn't help the qualia freak. Future events are in this respect just like present events. One day in the future, we will be in causal contact with the things we today claim to have knowledge about. According to the causal conception of knowledge, knowledge of the future (if there is such a thing) is parasitic upon normal causally connected types of knowledge.

II

The other possible limitation to a causal theory of knowledge was abstract objects. Abstract objects lack causal powers, but this need not make us think that there is anything particularly troublesome with acquiring knowledge about them.¹⁵ This would still leave the *perception* of qualia unaccounted for, but perhaps that can be dealt with in some other way.

But does anyone really believe that qualia are abstract objects, like sets and the natural numbers? Of course, they are abstract in the sense of not being physical, or that is what friends of qualia are prepared to tell us. But is this enough? Our very contact with them was

¹⁵ This has of course been debated, especially after Benacerraf's (1973). Physicalists have tended to see abstract objects' lack of causal powers as a factor that renders account of our knowledge of abstract objects problematic. See also the papers in Irvine (1989).

the essential thing about qualia, but our contact with sets and natural numbers is, it appears, not essential for these entities.

There is nothing qualitative about sets or natural numbers, but the very *raison d'être* of qualia is their qualitative nature, the way they feel. The (other?) abstract objects are in some loose sense theoretical constructs, posited for theoretical reasons, in order to account for something we have reason to desire an account for (*pace* Gödel's view of the abstract objects).

We could persevere, and try to introduce some new sort of perception of qualia, something relevantly like Gödel's conception of the intuition of mathematical objects. Ordinarily speaking, we do not perceive qualia, however; we may be said to have certain qualia, and the more usual understanding of our contact - if that is the correct word - with qualia makes this a special sort of cognitive contact.

III

But "contact", ordinarily understood, is again connected with a notion of causality. Perhaps one therefore wants to say that our contact with qualia is entirely *sui generis*, so that none of the usual notions of belief, perception &c apply, since the usual notions, it is admitted, are connected with causal notions. One wishes to rest with saying that we *have* qualia, not that we perceive them. Maybe this can be done, but the manœuvre has a price. We started out with the idea that qualia were peculiar entities, escaping the net of physical theory and causal interaction, and that their properties were not like the properties of ordinary objects. Now not even our connections with these strange entities are like our connections with other, ordinary, objects. But what does this leave of the notion of qualia? As far as I can see, there is little of interest left of the original conception of qualia. If this were all that can be said for, or about, qualia, there would be no need for Dennett to go through the steps of "quining" them.

IV

Yet another way to meet the objection is to partly give up the idea that qualia are epiphenomenal, and say that the instantiation of qualia has causal effects on other *mental* states.¹⁶ Given what we may want to say about other mental phenomena such as remembering, it is desirable that qualia should have some causal effects on other mental states, at least. (How else could one remember what a qualia was like? Why would one want to have that chocolate cake again?) This idea is still difficult to assess, since it appears to be *prima facie* plausible to hold that mental phenomena are connected, which here would mean that either all mental phenomena are epiphenomenal, or none are. Therefore epiphenomenal qualia lead to general epiphenomenalism, or else they are not really epiphenomenal.

It is obvious that more could be said about this, but this short paper was intended to discuss a *prima facie* conflict between the claim that qualia are epiphenomenal and a not unreasonable account of knowledge from a physicalist's perspective, so I will not discuss the ramifications of altering the conception of qualia with which we started.

V

Perhaps one response could be that Mary does not get to *know* anything new when stepping out of her confined room, but surely there is *something* new to her situation. Obviously there is something new for Mary - the whole description of the thought experiment was intended to get that much across. But is the difference relevant, or sufficient, for the demise of physicalism? Jackson supports his claim that Mary gets new knowledge by saying that her situation "is rightly described as *learning* - she will not say 'ho, hum'."¹⁷ Well, maybe she would not say, "ho, hum", or words to that effect, but what would this show? If the theoretical reasons against calling something knowledge are sufficiently strong, then there may be reasons for Mary to revise her first impression that she had learned something. If all that remained to be said about the thought experiment was that something happens, then such a

¹⁶Jackson suggests this himself on p. 474 of "Epiphenomenal Qualia", and refers the reader to his [1972].

¹⁷Jackson (1986), p. 392.

weak outcome can be discarded by the physicalist - there would for instance be room for an error theory of qualia.

VI

Does all of the above consist in question begging on behalf of the physicalist? Well, since the general idea was to delineate a possible strategy for the physicalist, it should be no surprise that certain assumptions that favour a physicalistic outlook are in place, but in at least one important aspect no questions are being begged. It is not assumed that the causal ties between the subject and the fact known need to be physicalistically acceptable, only that the general demand on knowledge is such that there must be *some* causally relevant tie. If there is such a thing as exclusively mental causation that could be referred to as well. The only claim is that qualia could not be genuine epiphenomena and objects of knowledge at the same time, and a further claim is that qualia cannot be epiphenomena without threatening to make all of the mental epiphenomenal in the process.

VII

Even if the physicalist has been given a new opportunity to withstand Jackson's thought experiment, something of the explanatory gap remains, along with a certain sense of mystery regarding what we should say about Mary's predicament. As I said in the introductory remarks, we have strong intuitions in favour of Jackson's construal of the thought experiment, and the physicalist manoeuvres that have been employed to make qualia safe for physicalism have only shown something pretty limited - that it is not reasonable to see them as epiphenomenal. In combination with a general physicalism, this means that qualia will be possible to handle by physicalists, but nothing has been said about *how* this is to be accomplished, and no one seems to have even the faintest idea about how events relating to consciousness and qualia are to be incorporated into physical theory.

It seems that we have a genuine dilemma here. Either qualia resist incorporation into physics, or they don't. If they do, then physicalism is false, but holding qualia to be epiphenomenal

leads to general epiphenomenalism. If they don't resist incorporation into physics, then *this incorporation* resists incorporation into our body of knowledge. (Not just our present state of knowledge, but also the foreseeable future.) This is more attractive than the other alternative, but it's still very unattractive - there is some reply, but we don't know which. Of course it could be that the reply is right around the corner, but the reasons for optimism are not overwhelming, at least not so far.¹⁸

References

- Benacerraf, P., "Mathematical Truth", *Journal of Philosophy* 70 (1973)
- Braddon-Mitchell, D., & F. Jackson, *Philosophy of Mind and Cognition*, Oxford 1996
- Bigelow and Pargetter, "Acquaintance With Qualia", *Theoria* 56 (1990)
- Churchland, P., "Reduction, Qualia and the Direct Inspection of Brain States", *Journal of Philosophy*, 82 (1985)
- Dennett, D., "Quining Qualia", in Lycan (1990)
- Heil, J., and A. Mele (eds.), *Mental Causation*, Oxford 1993
- Irvine, A. (ed.), *Physicalism in Mathematics*, Dordrecht 1989
- Jackson, F., Review of Campbell, *Body and Mind*, *Australasian Journal of Philosophy* 50 (1972)
- "Epiphenomenal Qualia", *Philosophical Quarterly*, 32 (1982) Reprinted in Lycan. Page references to that edition.
 - "What Mary Didn't Know", *Journal of Philosophy*, 83 (1986) Reprinted in Rosenthal. Page references to that edition.
- Lewis, D., "What Experience Teaches", in Lycan (1990)

¹⁸This pessimistic assessment is somewhat similar to the position presented in McGinn's (1991): "The deeper science probes into the brain the more remote it seems to get from consciousness. Greater knowledge of the brain thus destroys our illusions about the kinds of properties that might be discovered by travelling along this path. Advanced neurophysiological theory seems only to deepen the miracle." (p. 14n)

Lycan, W., (ed.), *Mind and Cognition: A Reader*, Oxford 1990

McGinn, C., *The Problem of Consciousness* Basil Blackwell, Oxford, 1991

Nemirow, L., "Physicalism and the Cognitive Role of Acquaintance", in Lycan (1990)

Papineau, D., "Physicalism, Consciousness and the Antipathetic Fallacy", *Australasian Journal of Philosophy* 71 (1993)

Pereboom, D., "Bats, Brain Scientists, and the Limitations of Introspection", *Philosophy and Phenomenological Research* 54 (1994)

Rosenthal, D., *The Nature of Mind*, Oxford U.P., Oxford 1991

Yablo, S., "Mental Causation", *The Philosophical Review*, 1992