

# Essays in Philosophy

A Biannual Journal

Vol. 4 No. 2, June 2003

## Book Review

---

*Knowledge, Possibility and Consciousness*, by John Perry. Cambridge, MA and London: The MIT Press, 2001. 221 + xvi pages (including index and notes). ISBN: 0-262-16199-0.

---

John Perry offers a spirited and sophisticated defense of physicalism against three well-known “neo-dualist” challenges, which are outlined in the first chapter. On stage are the zombie argument, the knowledge argument, and the modal argument. Having locked onto these primary targets, Perry devotes the next two chapters to outlining the basic tenets of his own view (which he calls “antecedent physicalism”), and in the remainder of the book he develops his responses to the opposition.

In contrast to older mind-brain identity theories, which identified all mental state types with physical types, Perry follows a more recent distinction between “psychological” and “phenomenal” aspects of mind. He favors a functionalist treatment of psychological states (e.g., beliefs) not essentially involving any subjective elements; the nature of psychological states is encapsulated by the roles that they play. However, phenomenal states (e.g., pains), which have essential subjective characters or “qualia”, require a different account, and here’s where Perry’s identity theory enters the picture: phenomenal states, he argues, are brain states. This type-identity claim carries with it a commitment to some degree of “neural chauvinism” – the view that only beings physically like us are capable of having the same phenomenal states that we enjoy – but Perry is willing to live with that. Anyway, as he notes, it doesn’t follow that extraterrestrials, robots, etc. could have no phenomenal states at all; rather theirs, if they had any, would just be different from ours.

Antecedent physicalism combines a number of commonsense ideas about experience with the claim that experiential states are (types of) physical states. Most important, phenomenal states are both causes and effects; they have subjective characters; they may have functional roles, but are not reducible to any such roles; and their nature is intrinsic, not determined by historical or contextual properties. Incidentally, the term “antecedent physicalism” should not be taken to suggest some sort of irreversibly entrenched dogma, but neither does it mark a position of ontological neutrality. Since neo-dualist arguments purport to show that physicalism is inadequate, Perry’s strategy is to start by embracing physicalism, and from there to assess whether neo-dualist considerations require it to be abandoned.

David Chalmers’ version of the zombie argument provides the first test. In a nutshell, it introduces a possible world, physically identical to ours – complete with exact duplicates of human bodies – but lacking qualia. On the supposition that a world physically identical to ours could lack phenomenal states, the zombie argument concludes that subjective experience must be nonphysical. Now, it seems obvious that antecedent physicalists should reject this reasoning in the following way:

subjective states are physical states; therefore a world physically identical to ours would contain phenomenal experiences; therefore the imagined zombie world is impossible. In short, from the perspective of antecedent physicalism, zombie scenarios beg the question.

Strangely, although Perry later endorses this point in passing, it is not the one that he actually employs against the possibility of zombie worlds. Instead, he focuses on antecedent physicalism's principle that phenomenal states are causally efficacious. He then proceeds to argue against the possibility of zombie worlds on the grounds that, absent phenomenal states, those worlds would lack the (physical) effects caused by phenomenal states in our world, and consequently no zombie world would be physically identical to the actual world. For example, assuming that the chocolatey and sugary qualia I get from biting into a Mrs. Field's cookie cause me to say, "That treat was delicious", a world in which my zombie counterpart ate such a cookie would fail to contain the same utterance (or if the zombie utterance did occur, causes other than qualia would be responsible). For this reason Perry concludes that a world lacking qualia would differ physically from ours, contrary to what the zombie argument supposes.

Why does Perry confront the zombie argument in such a convoluted way, given the availability of the more straightforward reply outlined two paragraphs back? The answer is that he thinks that zombie possibilities aren't fundamentally about physicalism vs. dualism. Rather, he holds that a commitment to epiphenomenalism would permit *either* dualists *or* physicalists to countenance the possibility of zombies. According to Perry, Chalmers' zombie world is impossible because "the antecedent physicalist believes in the efficacy of the conscious and rejects epiphenomenalism" (p. 72). But in order to show that physicalists and dualists are in the same boat here, Perry has to do some fancy footwork. First he introduces the notion of *physicalist* epiphenomenalists, for whom phenomenal states are "physical nomological danglers". Moreover, he acknowledges that physicalist epiphenomenalists cannot allow for the possibility of a zombie world differing from ours *only* with respect to the *effects* of conscious states, for if conscious states are physical states then any world physically identical to ours would have to include conscious states, whether or not they produced any effects. So, Perry further proposes that a physicalist epiphenomenalist "can accept the possibility of . . . zombie worlds . . . that are *physically indiscernible [from ours] except for the absence of the sensations*" (p. 78, italics added). Unfortunately there is trouble for Perry here, as the italicized phrase reveals. From the standpoint of physicalist epiphenomenalism, the re-characterized zombie world isn't Chalmers' zombie world at all. Once this point is recognized, there's no reason to suppose that epiphenomenalism is the central culprit in the zombie argument; Perry can't get that result by appealing to a zombie world that is physically distinguishable from the actual world in some respects. However, prior to introducing epiphenomenalist considerations, Perry already had all that he needed to disallow Chalmers' zombie scenario. Since antecedent physicalism maintains that phenomenal states are brain states, it follows that any possible world physically identical to ours will automatically preclude zombies, and this point is independent of any causal powers that phenomenal states might or might not have. Thus, it strikes me that antecedent physicalists should forget about epiphenomenalism and instead reject the zombie argument simply because it presupposes that experiential states are not brain states. In fact, that's the very strategy Perry does adopt in his reply to inverted-spectrum arguments.

Despite his dismissal of epiphenomenalism (in both physicalist and dualist versions), Perry at one point makes a surprising concession to that doctrine. After imagining a situation in which he picks

up a piece of red-hot charcoal, feels pain, and quickly drops the fiery coal, he writes:

It seems to me that the feeling of pain caused me to drop the charcoal. I may [be] wrong about that. It may well be that I drop the charcoal, quite independently of the feeling of pain; that the feeling of pain, and the release of the muscles that hold the charcoal, are both caused by more immediate effects of the heat of the charcoal on my nervous system, rather than the pain being the cause of the release, as it seems. There is no reason for the antecedent physicalist to think that we are always right about what conscious states cause.

(p. 76)

This is a perplexing comment from someone who, only two pages earlier, after describing the experience of biting into a chocolate chip cookie, found it “simply incredible” that the taste sensation didn’t cause him to voice his approval. It is hardly clear why the cookie example should weigh against epiphenomenalism if the charcoal example doesn’t.

Perry devotes chapters four through seven to dissecting Frank Jackson’s famous knowledge argument featuring Mary, the color scientist. Aficionados will recall that Mary, a scientific genius, has spent all of her life in a monochromatic environment where she learned all there is to know about the (completed) physics and neurophysiology of color perception. When she is released from her room and spies a ripe tomato, she discovers something that she did not previously know; namely, what it is like to experience red. Since prior to this moment Mary possessed all of the physical facts concerning color, and since she has now learned something new, the argument concludes that knowledge of what it is like to experience red cannot merely be knowledge of physical facts.

There is a structural parallel between Perry’s treatment of this argument and his earlier discussion of zombies. Once again he argues that the central issue is not dualism vs. physicalism, but rather an underlying assumption that both positions might adopt. This time, however, Perry’s view is more persuasive (for a caveat see the penultimate paragraph of this review). He contends that the knowledge argument rests on an erroneous conception of knowledge and belief. The mistake involves what Perry calls the “subject matter assumption” – that the *content* of a given belief is constituted entirely by truth conditions governing what the belief is about. This makes content independent of a representation system, and it raises a familiar puzzle about how identities can be informative.

So it is with Mary. While in her room, she learned that people have subjective experiences, including subjective experiences common to the perception of roses, fire engines, blood, etc. She hadn’t had this experience herself, and she didn’t know what it was like. Following Perry, call this subjective state “ $Q_R$ ”. When Mary was subsequently shown a ripe tomato, and exclaimed, “Oh, so *this* experience is  $Q_R$ ”, her phrase “*this* experience” denoted  $Q_R$ . But then it appears that Mary’s alleged discovery was no discovery at all, since Mary already knew that  $Q_R = Q_R$ . The knowledge argument concludes that  $Q_R$  cannot be a physical state, for Mary possessed all physical knowledge before seeing the tomato, and yet she has just learned something new.

Perry’s diagnosis of the situation contains many subtleties (canvassed in 75 pages), but the main

idea is this. What Mary learned cannot be described adequately by focusing on what's *true of* what she learned about. It's obviously *true of*  $Q_R$  that it is self-identical, and it's also true that when Mary first experienced red she was experiencing (something identical to)  $Q_R$ . On the subject matter assumption, the content of Mary's belief about  $Q_R$  is fixed by the facts about  $Q_R$ . The facts concerning  $Q_R$  didn't change when Mary first saw the tomato, and so it looks as if Mary didn't learn anything new about  $Q_R$ . Jackson takes this to show that the  $Q_R$  can't be a physical property, because Mary did learn something new. But Perry argues that physicalism is not the problem here. Rather, the difficulty is with the subject matter assumption, which insists that the content of Mary's belief about  $Q_R$  while in her monochromatic environment is the same as the content of her belief after seeing a red object. Perry contends that way of assigning content is seriously incomplete.

An example analogous to several of Perry's will serve to introduce the proposed repair. Suppose that you are on a backpacking trip in the wilderness (in pre-GPS days), and you become hopelessly lost. You suddenly spy a map on the ground that has probably fallen out of some previous hiker's pocket. Gratefully unfolding it, you see an area marked with an "X", and the notation "You are here". Unfortunately, though, this still doesn't give you any idea of where you are. However, suppose further that after wandering aimlessly for a while, you happen upon a kiosk erected by the Forest Service; it contains a map – perhaps even identical to the one you found earlier – also with an "X" and a corresponding legend. Unlike your previous discovery, this one is informative, but how so? After all, immediately prior to noticing the kiosk, your perceptions were of the very same place to which the new map's "X" refers, so there's only one location in question and it is obviously self-identical. But you already knew that everything is self-identical, so what exactly have you learned? As Perry emphasizes, your new knowledge cannot be described by concentrating exclusively on the referent of "here"; equally important is your newfound ability to *link* your perceptions of your surroundings with what the map says. You didn't learn that you were in a new place; rather, you gained a new way of identifying the very same place (first as "the place where I am now standing", and subsequently as "the place designated by the 'X' on the kiosk map"). Needless to say, this point has nothing to do with dualism vs. physicalism.

The wilderness example applies to Mary as follows. On Perry's analysis, the knowledge argument conflates (a) two different ways of knowing the same thing, and (b) knowing two different things. As was the case with you and the kiosk map, Mary didn't learn a new nonphysical fact; rather, she acquired a new way of knowing a previously mastered physical fact. She gained recognitional knowledge, expressible as " $Q_R$  is *this* subjective character", and her new knowledge was not the subject matter content that  $Q_R = Q_R$ . To capture the wanted notion, Perry develops an account of what he calls "reflexive content". This approach treats knowledge in terms of situated agents rather than simply as sets of propositions characterized independently of a representation scheme. Thus it includes, not just conditions on an object that is represented, but further conditions on an individual's representations, including the contexts in which they occur. In ignoring agents' perspectives, the subject matter assumption adheres to a false doctrine of objectivity, according to which "there is some kind of knowledge that involves grasping a fact not from any point of view – a view from nowhere" (p. 166). However as Perry aptly remarked earlier, "The view from nowhere is not a view at all" (p. 138). And antecedent physicalists are not committed to it. In rejecting the subject matter assumption, physicalists can also reject the claim that Mary could have known everything about a red experience prior to having one. Moreover, Perry argues that if dualists

adhered to the subject matter assumption, they too would have a major problem in explaining Mary's new knowledge.

Perry's final chapter addresses Saul Kripke's modal argument against the type identity of sensations and brain states (a similar one has also been used by Chalmers). Although mind-brain identity theses were traditionally put forth as contingent identities, Kripke has convincingly argued that identities are *necessary*. However, as Kripke also notes, it seems that one can coherently deny that sensations (e.g., pains) are brain states. If those denials point to a real possibility of nonidentity, then mind-brain identity doesn't hold for sensations, since the identity – if it holds at all – is necessary. Identity theorists might answer this by retreating to the claim that denials of identity here aren't really coherent after all, and any appearance of their being so is explainable as an illusion. Kripke considers various moves along that dimension, and argues that they fail. But if the relation between sensations and brain states really is contingent, there can be no identity.

In responding to Kripke, Perry grants that the mind-brain identity at issue is necessary, and he offers an alternative account of why it appears to be contingent. Perry has a number of interesting things to say about this, but the core idea again relies on the distinction between subject matter content and reflexive content, along with a corresponding distinction between what is possible and what is conceivable. At the subject matter level, experiential states are brain states, and there is no possibility that they aren't. But at the level of reflexive content, an agent's concept of subjective states may fail to link up with his concept of brain states even though they in fact denote the same thing, and in circumstances when the linkage fails it will be conceivable to him that there is no identity. Clearly, this point cannot be made simply by modeling the world(s) in which terms for various experiences and brain states denote the same thing, but reflexive content goes beyond that by modeling the way in which agents *represent* the world. As Perry puts it, "We need knowledge that reflects not just the way things are among themselves, but also how they are *for* us, how the ideas in our heads are connected to the subject matter they represent" (p. 176). It is at this epistemic level that the illusion of contingency finds its explanation.

*Knowledge, Possibility, and Consciousness* is a stimulating and lucid volume that breathes new life into a type-identity theory for qualia. Serious argumentation is leavened by frequent humorous lines, perhaps reflecting the book's origin as the 1999 Nicod Lectures. Perry's distinction between "subject matter" and "reflexive" content is obviously at the center of his defense of antecedent physicalism, and it is tempting to wonder whether the latter might be in principle reducible to the former. Perry's response to any such suggestion is clearly negative, and he may well be right, but then additional perplexities naturally arise. For example, Perry allows (1) that Mary did acquire a new belief, and he argues (2) that her new belief cannot be accounted for in terms of subject matter content. In order to accommodate both claims, he counts new reflexive content as new belief content. Although Perry's use of reflexive content here is undeniably significant and powerful, he acknowledges that its introduction requires jettisoning standard accounts of "what is believed". But do new reflexive contents really yield new beliefs or new knowledge? This and surrounding matters would benefit from further exploration. Finally, it should be mentioned that the index to the book is rather sparse; there are entries for "Raquel Welch", "Reno, Nevada", and "orgasm", but none for such important specialized notions as "concepts", "notions", and "perceptual buffers".

I am grateful to Candace Bolter and Nathan Oaklander for helpful comments on an earlier draft of this review.

Charles E. M. Dunlop  
The University of Michigan-Flint

---

Copyright © 2003, Humboldt State University