

Book reviews

Knowledge, Possibility, and Consciousness, by John Perry. Cambridge, MA: MIT Press, 2001, xvi + 221 pp, \$18.00.

The hard kernel of the problem of consciousness is the fact that the phenomenal consciousness with which we are acquainted in our own case is quite different from anything that can be observed in the brain. As John Perry quotes the philosopher A.C. Ewing,

We know by experience what feeling pain is like and we know by experience what the physiological reactions to it are, and the two are totally unlike ... the difference is as plainly marked and as much an empirical matter as that between a sight and a sound (p. 3, quoting A.C. Ewing, *The Fundamental Questions of Philosophy*, New York: Collier Books, 1962, p. 100).

How could two things that seem so different in fact be the same? Perry thinks that the physicalist must concede that there are two different ways of knowing: that our awareness of our own experiences is epistemically quite different from our access to brain states (through perceiving them or reading descriptions of them), with two correspondingly different associated vocabularies. The question is whether that means that there are two things known about. The seminal identity theorists of the mid-twentieth century, such as Place, Smart and Feigl, argued that, although the mind-brain identity thesis might be a surprising one it could nevertheless be true, on the model of other surprising-but-true identities such as lightning being atmospheric electrical discharge or water being H₂O. Perry diagnoses this as a 'semantical turn' in the philosophy of mind, a shift that forged a connection between the problem of consciousness and Frege's problem: how is it that a true identity statement — that is, the statement that something is the same thing as itself — can be informative?

Perry sees this semantical turn as crucial to a proper defence of physicalism, and his interesting and instructive book is an extended pursuit of its lessons. The solution to the problem of consciousness, according to Perry, thus turns out to consist in a solution to a version of Frege's problem: the physicalist needs to explain why the physical

and the phenomenal *seem* so different in a way that is consistent with them in fact being the same thing.

The strategy founded by J.J.C. Smart was to appeal to topic-neutral concepts of properties: we are to identify the property the brain state has that makes it an experience — that gives it the apparently special first-person mode of presentation of a conscious state — by applying a concept that is neutral between the property so-characterised being physical or non-physical. One important descendent of that strategy ‘has been to find this neutrality in causal-role or functional properties that a brain state might have’ (p. 205) but Perry argues that this method fails for phenomenality. ‘We need instead the topic-neutrality of demonstrative/recognition concepts’ (p. 205). The neutral concept in question will be something like ‘this property’ or ‘the property currently being attended to by Mary.’ That *this* sensation is brain state *B* is an informative identity, since it involves two quite different modes of presentation (one neutral and one physical) of the same thing.

So far, all of this might seem quite familiar to aficionados of the work of, say, Brian Loar and Bill Lycan. The main thesis of Perry’s book, though, and his main novel contribution to the consciousness debate, is that participants on both sides have been labouring under what he calls ‘the subject matter fallacy.’ The modes of presentation that distinguish the two ways of knowing about conscious states cannot be understood in terms of attributive conditions of reference — in terms of different subject matters. For to do so, would be to introduce a second property, in addition to the physical one, to fill the role of the mode of presentation of experience. Instead:

My solution has been to explain the twoness, in its various forms, not at the level of what is known about, but at the level of what is involved in the knowing: the level of reflexive content (p. 207).

What Perry means by ‘reflexive content’ is best explained by utilising what he calls his ‘content analyzer’: Given *such and such*, φ is true iff *so and so*. Where φ is a truth-evaluable representation and *such and such* are *facts about the representation*, then a set of truth conditions are generated for φ . Thus, by varying what is given we vary the content assigned.

These will not be different theories about *the* content of φ . They will be ways of getting at different systematically related contents of φ (p. 125).

Our ordinary concept of content is ‘referential content,’ which is generated by fixing all the facts about the representation, including the meanings and referents of the words and the context of their utterance. But if we do not fix all of these facts about φ — if we allow them to vary — then we add them, as it were, to the truth conditions of φ ; we output a richer *content* for φ . One of Perry’s examples is the utterance (1) ‘That man [pointing to Dretske] wrote *Knowledge and the Flow of Information*.’ The referential content of (1) is something like (1’) ‘Dretske wrote *Knowledge and the Flow of Information*.’ But, as Perry points out, I might easily know (1’) without knowing (1) — to come to know (1) might well be to *increase* my knowledge, even if I already knew (1’). How can this be, if (1) and (1’) have the same referential content? Perry’s answer is that (1) has more content available if we do not ‘load’ the facts about context: then we get something like (1’’) ‘The person the speaker of (1) is ostending is the author of *Knowledge and the Flow of Information*.’

A crucial thing to notice about (1’), for Perry, is that *this* content is not any part of the *subject matter* of (1). (1) itself — which appears as part of the specification of (1’’) — is not part of its own subject matter; we do not, as a general rule, say things about our own utterances. The truth conditions that the content analyzer revealed here are what Perry calls *reflexive* truth conditions: conditions (partly) on the representation itself, not just on the subject matter of the representation.

The point Perry makes here is, of course, not particularly about phenomenal consciousness, but is a general semantic phenomenon. This is how his work in this book connects with his well-known previous work: his view is that reflexive content is crucial to understanding not just indexicals but the ways in which thought connects with perception and action. But it is a general point that has important implications for physicalism, Perry urges, and neglect of it has been responsible for the apparent plausibility of neo-dualism in recent years.

First, Perry formulates what he calls ‘antecedent physicalism’: not that physicalism is demonstrably true but that anyone who antecedently accepts the main tenets of commonsense physicalism — for Perry, that is basically mind-brain identity theory in the tradition of Herbert Feigl — has the resources to repel the attacks of the neo-dualists. He identifies three primary arguments that are relied upon to bolster the neo-dualist case — the zombie argument, the knowledge argument, and the modal argument — and throughout the book he attempts to refute each in turn.

He disposes of the zombie argument, initially, in quite short order. The zombie argument runs as follows: there is a possible world whose inhabitants are physically indiscernible from us but are not conscious; hence qualia do not logically supervene upon our physical states (since it is possible for them to come apart from that physics); and so consciousness must be something over and above the physical. Perry's response is that the antecedent physicalist need not accept the first premise here: if indeed states of consciousness are *identical with* (rather than merely supervenient on) the physical, then it must be impossible to remove consciousness and leave the physical exactly the same. This is especially the case because the antecedent physicalist holds that conscious states are causally efficacious, and so if removed from the web of causation would change the way things go in the world. This does not establish that physicalism is true, but it is enough, Perry says, to show that the antecedent physicalist need not be persuaded by the zombie argument.

The knowledge and the modal arguments are addressed by appealing to the notion of reflexive content. The knowledge argument says that someone (say Mary, a brilliant scientist raised all her life in a black and white environment) could know *all* the physical facts about colour, and yet learn something quite new when they see colour for the first time; thus, there must be more facts than physical facts, so physicalism is false. Perry's response is that Mary does indeed learn something new, and what she learns is indeed knowledge — her belief states change in a way that changes their truth conditions — but she does not come to this new knowledge by learning a new *fact*. The assumption that she does depends upon 'a distorted picture of the relation between knowledge and reality, between epistemology and metaphysics' (p. 16). Mary thinks 'This is what it is like to see red' and part of the content of this thought is the new thing that she has learnt; but it is reflexive content, not subject matter content, that is new. The mode of presentation that distinguishes Mary's introspective access to colour experience from her scientific access does not involve a new property — a previously unknown aspect of the brain state — but has to do with the new way that her representations *link up* with the world and each other (just as I might learn that *this* is Fred Dretske by having my pre-existing Dretske concept link up my present perceptual content in a way that it was not linked before).

The modal argument works by noting that we have a very strong sense that the connection between the phenomenal and the physical is

contingent: the brain state associated with pain feels like *this*, but it might easily have felt like *that* (and so not been painful) or the painful feel might instead have been connected to some other brain state (activity in c-fibres instead of pyramidal cells, for example). Yet, as Kripke most forcefully pointed out, if ‘two things’ are identical then they are *necessarily* identical — nothing could be anything other than itself. In some cases, the appearance of contingency can be explained away and an identity established; typically, this is done by showing that we are confusing the necessary identity with the contingent fact that some object fits the particular identifying criteria associated with one of the concepts involved. But there seems to be no room to explain away the contingency in this manner in the case of pain: having *this* feeling *is* what it is to be in pain, and is not a merely contingent way of identifying the pain, and the problem precisely is that *this* feeling seems able to come apart from the brain state.

Perry’s rebuttal of this argument rests, again, on the notion of reflexive content. At the level of subject matter content — that is, the properties and things in the world the possibilities are possibilities for, after we have fixed all the facts about the denotations of our concepts — there might be no possibility that pain is not pyramidal cell activity. The appearance of contingency, Perry argues, comes from the various possibilities for the way that thought and language might link up with the world: my phenomenal concept of pain might have been about something other than pyramidal cell activity, and when I learn that it is not, that possibility is eliminated — I learn something new. But what I learn is not a new extra-linguistic fact; it is not a new piece of subject matter (my thought about pain being a brain state is not a thought about my pain concept); it is a new way in which my thoughts are linked together and to the world. (Here Perry circles back to the zombie argument: the deeper reason why zombie arguments fail, he suggests, is that zombie possibilities are not genuine possibilities but merely linguistic possibilities about how my concepts might have referred, and so on.)

So the solution that Perry proposes to the problem of consciousness is a logico-semantic one; however, I think there are a series of inter-connected problems about Perry’s notion of reflexive content that draw this contention into doubt, and reveal the deeper metaphysical issue that lies untouched below Perry’s discussions. Here, in brief, are three of the problems. First, Perry seems to hold that all beliefs and utterances have reflexive as well as subject matter content:

after all, there are always facts about denotation and context that can either remain fixed or not. But identity statements *vary* in the degree to which they seem contingent: some are *a priori* necessary, some necessary only *a posteriori*, and some — such as psychophysical identities — do not seem necessary at all. For example, Perry insists (p. 143–4), mathematical identities can be non-obvious because of their reflexive content. But mathematical identities, unlike psychophysical ones, are necessary *a priori*. So it is not clear that appeal to reflexive content is by itself enough to explain the differences between the informativeness of these identities, even though it is just this that Perry has set out to explain.

Second, as several critics have noted, there is a powerful *prima facie* disanalogy between the cases of perspectival belief that Perry uses to ground his position, and phenomenal belief. Typically, in indexical or otherwise perspectival cases, once the link has been made to a non-perspectival version of the representation it is clear that the subject matter contents of the two representations are the same. For example, when I realize that the person before me is Dretske, then it is easy to see that my perception and the name ‘Dretske’ denote one and the same thing — a person of such-and-such physical features, with such-and-such a history, and so on; the concepts are relevantly isomorphic and overlapping. By contrast, even if we were to accept that as a matter of fact phenomenal and physiological concepts *might* co-refer, it would remain far from clear *that* they do so.

Finally, consider the case of non-linguistic animals. Perry’s account is explicitly linguistic, and he concedes that therefore, in his example, a mouse caught in a trap is not conscious of its pain (‘Replies,’ in *Philosophy and Phenomenological Research* 68, 2004, p. 225). Yet he does think that there is *something it is like* to be that mouse, which has nothing to do with reflexive content; for these ‘what it is like’ properties, ‘I have no account of them at all. I acknowledge that there are such states, and I argue that we have been given no good reason to suppose they are not physical’ (‘Replies,’ p. 225). ‘What it is like’ to be in pain is surely at the heart of the problem of consciousness; it is disappointing to discover that Perry’s stimulating and rich book nevertheless ducks this question.

Andrew Bailey (abailey@uoguelph.ca)
Dept. of Philosophy, University of Guelph
Guelph, Ontario, Canada N1G 2W1