Theoria, 72, (1), 2006, 80-6.

David Papineau, *Thinking about Consciousness*. Oxford: Clarendon Press, 2002. 266 pages.

In this book, David Papineau sets two main tasks for himself. The first is to defend a materialist view about consciousness. The second is to explain why materialism should seem so counterintuitive and unsatisfactory, even to some of its proponents. Chapter 1 and a lengthy appendix are devoted to the first task. The main body of the book, chapters 2-6, is devoted to the second. These chapters include treatments of familiar challenges to materialism, such as the knowledge argument, Kripke's modal argument, and the explanatory gap. In the final chapter 7, Papineau argues that there are principled limits to making substantial scientific progress on consciousness. This review will be limited to the two main tasks of the book: the defence of materialism, and the diagnosis of why this view seems, to many, unsatisfactory.

Papineau defends a property-identity materialism: conscious properties are identical with material properties, where 'material property' covers both the first-order physical properties that are studied by the physical sciences and "higher-order", physically realised properties like functional and disjunctive properties. There is, he thinks, "one definitive argument" for this view (p. 15). This is "the causal argument", which in Papineau's preferred version rests on the following premises (pp. 17-19).

(1) Conscious facts have physical effects.
(2) All physical effects are fully caused by purely *physical* prior facts.
(3) The physical effects of conscious facts are not always overdetermined by distinct causes.

Papineau takes property-identity materialism to follow from these premises. By (1) and (2), there are effects that have both conscious and physical causes. And by (3), these effects don't have distinct causes. The "only possibility left" is that the conscious causes mentioned in (1) are identical with some part of the physical causes mentioned in (2) (p. 18). And that conclusion in turn trivially entails the property-identity claim, because facts are identical only when the properties they involve are identical (p. 19).

As Papineau observes, this version of the argument makes the controversial assumption that causes are *facts*, rather than particulars, like events. But he thinks a parallel argument for property-identity materialism is available to those who take causes to be events. In the parallel argument, this view is inferred from the following premises: (1)´ conscious events have physical effects *in virtue of* their conscious properties; (2)´ all physical effects are fully caused by prior physical events *in virtue of* the latter's physical properties; and (3)´ the physical effects of conscious events are not always caused *in virtue of* distinct properties of their prior circumstances (p. 20).

This revised argument actually relies on another assumption, namely that events (supposing they are causes) cause other events *in virtue of* some rather than other of their properties. Papineau doesn't mention that this assumption has been questioned as well (see Davidson 1993), and he doesn't defend it. So, his defence of materialism is not without gaps, but I shall not pause to pursue this issue.

The causal argument is familiar in contemporary philosophy of mind, but Papineau makes an important contribution to its defence. It is sometimes suggested that materialism is a philosophical fashion, which may seem credible because of its association with the prestigious physical sciences, but which in fact rests on metaphysical speculation unrelated to scientific results (for something like this view, see Burge 1993). Papineau takes issue with this view. According to him, the premises of the causal argument are justified, not by intuition or apriori reasoning, but rather by the standard considerations governing theory choice in science, such as simplicity and consistency

with data. Moreover, an important part of the data relevant to the causal argument has become available only by relatively recent empirical research.

His most extensive discussion on this topic concerns premise (2), the thesis that physics is causally complete. As Papineau notes, this premise, while widely believed today, was far from universally accepted prior to the twentieth century. But its new status is due, he argues, not to changes in intellectual fashion, but rather to the gradual accumulation of scientific evidence. In a historical overview ranging over Descartes, Leibniz, Newton, Helmholtz and others, Papineau tries to show that prior to the twentieth century, one couldn't rationally rule out the possibility that irreducibly vital or mental forces operated inside animate bodies. The important development in the twentieth century was that physiology and biochemistry began to explore the inside bodies, down to the level of the cell. All evidence from this research indicates that the forces operating inside bodies reduce to forces that operate elsewhere in nature. The completeness of physics is thus by now "a premise which, by any normal inductive standards, has been fully established by over a century of empirical research" (p. 256).

This "naturalistic" defence of the causal argument is, it seems to me, a significant contribution to the case for materialism. It may not be the most heavily emphasised part of this book, but to my mind it is the most convincing part.

The core of the book is Papineau's explanation of why materialism seems, to many, so unsatisfactory. The problem is the following. Normally, once we are given convincing enough evidence for an identity claim, any puzzlement we might have had about it evaporates. Take, for example, the claim that Superman is Clark Kent. Until we have sufficient reason to believe in this claim, we may wonder how it could possibly be true. How could Superman, the flying hero, possibly be identical with Clark Kent, the mild-mannered reporter? However, once we are provided with enough evidence that the identity claim is true, no sense of puzzlement remains. In the case of consciousness and the brain, on the other hand, the situation is different. Many philosophers are by now convinced -- often by the causal argument or some variety thereof -- that consciousness

must be identical with some property of the brain. But, to many, this identity *continues* to seem puzzling even after the evidence for it has been absorbed.

One possible diagnosis of this is that we presently lack the theoretical sophistication to really understand the materialist's identity claim. This is Thomas Nagel's view. He suggests that the present status of materialism is "similar to that which the hypothesis that matter is energy would have had if uttered by a pre-Socratic philosopher" (1974, p. 447). Just as the pre-Socratic philosopher lacked the theoretical background to fully grasp how matter could be energy, we lack the background required to really understand how consciousness could be identical with some aspect of the brain.

Papineau's diagnosis is different. On his view, the problem is not that we lack some theoretical insight required to understand the relevant identity claim. We know what conscious properties are, what brain properties are, and what it is or would be for them to be identical. The problem about materialism is that some of the concept we use to think about conscious properties generate an *illusion* that these properties must be distinct from brain properties. Because of this illusion, we can't really convince ourselves that materialism is true: "something stops us *really* believing the materialist identification of mind with brain, even those of us who officially profess materialism" (p. 94). This situation calls, not for better theories, but for a kind of philosophical "therapy" (p. 4). We need to identify the source of our instinctive resistance to materialism in order to free ourselves from it.

This general idea -- that our concepts for thinking about consciousness generate an illusion of mind-brain distinctness -- has been explored before, notably by Loar (1990), Lycan (1996, sect. 3.4), Tye (1999) and Papineau himself (1993a), (1993b, chapt. 4) and (1995). But this book is probably the most detailed and worked-out version of this idea to date.

According to Papineau, we can think about conscious properties both under "phenomenal" and under "material" concepts. To think about conscious properties under

phenomenal concepts is to think about them "in terms of *what they are like*", whereas to think about them under material concepts is to think about them "as items in the third-personal, causal world" (p. 48).

Phenomenal concepts are, on Papineau's view, a peculiar species of concept in that they "incorporate" versions of the very things they refer to (p. 117). To illustrate, think phenomenally about a pain you once felt. According to Papineau, when you do this, you imaginatively re-create a version of the original experience, and the re-created experience is part of the "term" you use to refer to the experience in question. This kind of reference is analogous to our practise of referring to linguistic signs by quoting the signs themselves. Thus, consider: the name 'David'. I just referred to a certain name, and I did so by quoting that name. Similarly, to think phenomenally about a certain experience is to think: the experience:---, where a version of the experience itself is plugged into the open slot.

> [O]rdinary quotation marks can be viewed as forming a frame, which, when filled by a word, yields a term for that word. Similarly, my phenomenal concepts involve a frame, which I have represented as 'the experience:---'; and, when this frame is filled by an experience, the whole then refers to that experience (p. 117).

This "quotational model" for phenomenal concepts explains our persistent intuition of mind-brain distinctness, according to Papineau. We resist materialism because *thinking* about experiences materially is so different from *thinking* about them phenomenally. When we think about experiences materially, we don't activate versions of these experiences. And since we don't, there is indeed a sense in which experiences are -- as it is often put -- "left out" when thought about materially. They are left out in that they are not *used* to think about themselves. Now, it is, Papineau claims, "terribly natural" to slide from this to the conclusion that material concepts don't *refer* to experiences (p. 170). The inference is, of course, fallacious. Material concepts may well *refer* to experiences despite not *using* them, and the causal argument establishes that some of them do. But it

is this fallacy, dubbed "the antipathetic fallacy" (p. 171), that accounts for our residual dissatisfaction with materialism.

The account is elegant. However, I don't think it is successful. There are, I think, at least two severe problems with it. First, *(i)* I doubt that phenomenal concepts work the way that Papineau suggests. Second, *(ii)* even if they worked that way, I don't see that this by itself would explain the intuition of mind-brain distinctness. I will develop these points in turn.

(i) I accept that *sometimes* when we think about a given experience, we imaginatively create a version of that experience. But I don't think such activation of experiences is essential or necessary to thinking -- even thinking phenomenally -- about experience. I can think about a pain in an entirely "casual" manner, without dwelling on or imaginatively conjuring up how it feels. And it seems to me that such a thought may well be phenomenal, in Papineau's sense.

To think phenomenally about a conscious property is, to repeat, to think about it "in terms of what it is like" rather than "as an item in the third-personal causal world". As Papineau explains, thinking about a conscious property in terms of what it is like is to think about it "directly" and not via some contingent property of it: "There is only one property in play when a phenomenal concept refers to a phenomenal property: namely, the phenomenal property itself. No further property mediates between referring concept and referent" (p. 104). Now, it seems to me that a "casually" entertained thought about pain could have exactly these characteristics. If I'm sufficiently unreflective and uneducated, it may never have crossed my mind that a pain *could* be an item in the third-personal, causal world. If so, I would not -- at least not knowingly -- *ever* think about a pain as such an item. Moreover, it seems that a casually entertained thought may well refer "directly" to its referent. When I think casually about a pain, there may be no property "in play" other than the pain itself. Notice also that there may be no doubt in my mind that a casually entertained thought about a pain co-refers with a thought that I entertain imaginatively on another occasion. In fact, it may be transparently and

immediately clear to me that these thoughts co-refer. Therefore, it is doubtful that any "cognitive significance"-test for distinctness of thoughts could give us reason to distinguish them.

(ii) Suppose, however, that the "quotational model" for phenomenal thinking about consciousness were correct. That would explain the intuition of mind-brain distinctness if it were, as Papineau holds, "terribly natural" to infer that material concepts don't *refer* to experiences given that they don't *use* experiences. But would this really be "terribly natural"? I don't see that it would, or that Papineau has given us reason to think so. One test relevant to this question is to look at the case of ordinary quotation, which by Papineau's own lights is the closest analogue to the working of phenomenal concepts. Thus, consider the concepts *the name 'David'* and *the first name of the author of Thinking about Consciousness*. The former concept uses the name 'David' to refer to that very name; the latter does not. But surely, this fact does not give rise to any inclination to think that these concepts can't co-refer. Once I absorb satisfactory evidence for their co-reference, no puzzle remains. Hence, it seems it can't be merely the fact that phenomenal concepts use versions of their referents (supposing that were in general true) that explains the persistent intuition that they can't co-refer with any material concept. But then Papineau's account of the intuition of mind-brain distinctness is at best incomplete.

So I don't think Papineau succeeds in putting the mind-body problem to rest in this book. But the book is a most rewarding reading despite this. It provides an unusually detailed and lucid articulation and defence of one approach to the mind-body problem. As such, it greatly contributes to clarifying our options in this area of philosophy.[1]

References

Burge, Tyler (1993) "Mind-body causation and explanatory practise", in Heil and Mele (1993).

Davidson, Donald (1993) "Thinking causes", in Heil and Mele (1993).

Heil, John and Alfred Mele (eds.) (1993) *Mental Causation*, Oxford: Clarendon.

[1] Thanks to Sten Lindström for pointing out an important unclarity in an earlier version of this text.

Loar, Brian (1990) "Phenomenal states", in *Philosophical Perspectives* 4*: Action Theory and Philosophy of Mind*, ed. James Tomberlin, Atascadero, CA: Ridgeview. Reprinted in revised form in Ned Block et al. (eds.) (1997) *The Nature of Consciousness: Philosophical Debates*, Cambridge, Mass.: MIT Press.

Lycan, William (1996) *Consciousness and Experience*, Cambridge, Mass.: MIT Press.

Nagel, Thomas (1974) "What is it like to be a bat?", *Philosophical Review* 83: 435-50.

Papineau, David (1993a) "Consciousness, physicalism and the antipathetic fallacy", *Australasian Journal of Philosophy* 71: 169-183.

Papineau, David (1993b) *Philosophical Naturalism*, Oxford: Blackwell.

Papineau, David (1995) "The antipathetic fallacy and the boundaries of consciousness", in *Conscious Experience*, ed. Thomas Metzinger, Paderborn: Schöningh-Verlag.

Tye, Michael (1999) "Phenomenal consciousness: The explanatory gap as cognitive illusion", *Mind* 108: 705-725. Reprinted as chapter 2 of Tye (2000) *Consciousness, Color, and Content*, Cambridge, Mass.: MIT Press.

Pär Sundström
Department of Philosophy and Linguistics
Umeå University
Email: *par.sundstrom@philos.umu.se*