

Could There Be A Science of Consciousness?
David Papineau

I Introduction

In this paper I want to consider the implications of materialism about the human mind for a scientific understanding of consciousness. I shall argue that, while science can tell us many exciting things about human consciousness, it won't be able to pinpoint any specific material property that constitutes seeing something red, say, or being in pain, or indeed that constitutes being conscious (that is, feeling like something rather than nothing). Not that this means there are definite facts about consciousness about which science must permanently remain silent. Rather the difficulty lies with our concepts of conscious properties, which are vague in certain crucial respects.

II Background Assumptions

Let me begin with four assumptions which will provide the framework for my argument, but which it will not be possible to defend properly here. (For detailed arguments, I refer readers to Papineau, 2002.)

(1) Materialism. Conscious properties are material properties.

Comments:

By 'material' properties I mean either strictly physical properties or properties which metaphysically supervene on strictly physical properties.

Materialism is strongly supported by the causal completeness of physics, which leaves no room for anything which is metaphysically independent of the physical realm to have any physical effects. If conscious properties were non-material, they would thus be epiphenomenal 'danglers', caused by physical occurrences, but themselves having no effects on physical activities. In this respect, they would be quite unique: nature offers no other examples of such impotent epiphenomenal properties.

Of course, if there were compelling independent grounds for holding that conscious properties are non-material, then we would have no option but to accept epiphenomenalism about consciousness. However, I do not think that the standard anti-materialist arguments (Kripke, 1980, Jackson, 1986) provide any such compelling grounds. So I see no reason to reject the relative simplicity of materialism for the metaphysical oddity of epiphenomenalist dualism.

(2) Conceptual Dualism. Nevertheless we have two quite different kinds of concepts for thinking about conscious/material properties, which I will call 'phenomenal' and 'material' concepts.

Comments:

'Phenomenal' concepts are most easily thought of as akin to demonstratives, or even quotational terms: we use a mental construction which can be represented as 'that experience: ---'; we fill the gap with some current experience or its recreation in imagination; and we thereby form a term which refers to that experience itself (either the type or the token). Compare the way that the term 'cat' refers to the word enclosed in the quotation marks (either the type or the token).

I shall use 'material concepts' as a catch-all for all other, non-phenomenal ways of thinking about conscious/material properties, including concepts of brain states, concepts of causal roles, and so on.

There are no a priori connections between phenomenal and material concepts. It is always informative to learn that that some given experience: --- (and here we 'quote' some experience) is constituted by some physical property, or plays a certain causal role, or has some other such material feature.

I take there to be a phenomenal concept of the determinable property conscious (being like something rather than nothing), as well of determinates like seeing something red or being in pain. This is why it always makes a priori sense to ask, given any material specification of some property, whether creatures with that property are phenomenally conscious. (For more on the structure of this determinable phenomenal concept of consciousness-as-such, see Papineau, 2002, pp.185-6.)

(3) Hard Problem Research. Consciousness research aims to identify the material referents of phenomenal concepts.

Comments:

David Chalmers (1996) distinguishes the ‘hard problem’ of figuring out what phenomenal concepts refer to from the ‘easy problem’ of showing how psychological concepts of causal roles are physiologically realized. While I of course disagree with Chalmers’ eventual conclusion that phenomenal concepts refer to non-material properties, I agree that it is the hard problem of identifying the referents of phenomenal concepts that makes consciousness research seem so exciting.

Thus there would be no special challenge about understanding colour vision if all we wanted to know were which physiological mechanisms realize the psychological concept of responding differentially to coloured objects. By contrast, the kind of consciousness research I shall be concerned with raises such questions as which property constitutes seeing something red—and here we mentally ‘quote’ the experience.

(4) Subjects’ Reports. Consciousness research hinges crucially on subjects’ phenomenal reports about present and past experiences.

Comments:

I take subjects’ first-personal phenomenal reports to provide a kind of observational data-base for research into the material referents of phenomenal concepts. Compare the way in which research into the material nature of water, say, must begin with observationally identified samples of water.

I shall say more about the status of subjects’ reports later. But we can bring out the importance of such reports by considering research into creatures who cannot make them, like vervet monkeys, say. You can stimulate vervet monkeys, you can get them to perform various tasks, and you can check on what is going on in their brains. But since the monkeys can’t you tell what they consciously experience at different times, none of this will cast any immediate light on the monkeys’ phenomenal consciousness. For all this research establishes, the monkeys might share very similar experiences to our own, or have no consciousness at all, or be somewhere in between. By contrast, humans can tell us when they are seeing something red, feeling a pain, or experiencing something rather than nothing, which offers a crucial handle by which to identify the material natures of phenomenal properties.

III Empirical Research into Consciousness

Given the points just made about subjects’ reports, it may seem that research into the material referents of phenomenal concepts presents a standard scientific problem. Suppose that we have a data-base of cases for some phenomenal property \emptyset , derived from subjects’ reports. We can then seek some material property M which is present whenever subjects report \emptyset , and is absent whenever subjects deny \emptyset . This strategy thus promises to identify M as the material nature of \emptyset , as desired.

I shall argue, however, that such research will always be less than fully effective, and will inevitably fail to pinpoint a unique material referent for any phenomenal concept. The is because there will always be a numbers of different material properties M which fit the evidence of the data base equally well, and between which the standard research methodology is therefore impotent to decide.

Of course, scientific research can do a great deal to narrow down the possible material referents of phenomenal concepts. If some material M is to be identical with phenomenal \emptyset , then it cannot be absent when \emptyset is testified to be present, otherwise it won’t be necessary for \emptyset . Nor can it be present when \emptyset is testified to be absent, otherwise it won’t be sufficient for \emptyset . Scientific research can show that many candidate Ms violate one or the other of these requirements. Take the phenomenal property of being in pain, say. Any material property that is sometimes absent when subjects report being in pain—some very specific neuronal activity that is peculiar to stabbing pains, for example—is thereby disqualified as unnecessary for pain as such. And any material property that is sometimes present when subjects say they are not in pain—some general neuronal activity common to both pain and heat sensations, say—is thereby disqualified as insufficient for pain.

Research of this kind can tell us many interesting and indeed surprising things about consciousness. For example, recent investigations have shown that subjects report no conscious experience of many cognitive processes which we might pre-theoretically have expected to be conscious: these include the visual guidance of hand movements (Weiskrantz, 1986, Goodale and Milner, 1992) and the instigation of voluntary actions (Libet, 1993). There are also cases of the converse kind, where subjects report conscious experiences in cases where we might have expected none: for example, subjects who take morphine to allay an existing pain will report that they can still consciously feel the pain, even after it has ceased to agitate them (Dennett, 1978).

Yet science can only take us so far. It may be able to rule out certain candidates as the referents of phenomenal concepts, often surprisingly so. But the trouble is that, even after science has narrowed things down in this way, there will always remain a plurality of Ms in play for any given phenomenal \emptyset , and no further means of deciding which is the real referent of the phenomenal concept.

My eventual diagnosis of this situation will not be that there are definite facts about consciousness to which we lack epistemological access—that there is some material property that really constitutes being in pain, say, but which we can't find out about. Rather, I shall argue that our phenomenal concepts of conscious states are vague—nothing in the semantic constitution of phenomenal concepts determines precisely which of the candidate material properties they refer to.

IV Too Many Candidates

To illustrate the basic difficulty, suppose that we have identified some strictly physical property P that is present whenever humans testify to being in pain, and absent whenever they deny this. If this is so, then will also surely be a number of more structural properties which are similarly perfectly correlated with human reports of being in pain. Simply focus on some detailed level of causal organization ensured by P, but abstract away from the specific physical way that this organization is realized in human beings. The resulting structural property S will also be present whenever humans report being in pain, and absent whenever they deny it. So there will be no way of using standard scientific research into humans to decide whether P or S constitutes being in pain.

To make the point graphic, consider a 'silicon doppelganger' who shares some detailed structural property S which co-varies with pain in humans, but whose basic chemistry is silicon-based rather than carbon-based. If you prick this being with a pin, you will produce S, but not the strictly physical P which also goes with pain in humans. So standard consciousness research leaves us with an obvious but unanswered question. When you prick the doppelganger, does it feel pain? If pain = S, then it does. If pain = P, then it doesn't. And if standard consciousness research on humans can't decide between S and P as the referent of being in pain, then it can't tell us what the doppelganger feels.

Note that there won't only be one alternative structural S here, but a whole sheaf of them. We can characterize humans physiologically, neuronally, and computationally, as well as strictly physically (cf. Lycan, 1987), and we are likely to find Ss at all these further levels that correlate perfectly in humans with any strictly physical P that goes with pain in humans. Empirical research will therefore be unable to decide between the strictly physical P and any of these candidate Ss as the referent of any given phenomenal concept.

It gets worse. As I shall show in section VII below, there is also a plurality of material candidates in the 'broad-narrow' dimension, as well as in the physical-structural dimension. Still, let us keep things simple for the moment, and stick with a simple pair of candidates, a strictly physical P found in humans, and some structural S which is common both to humans and their silicon doppelgangers.

IV More Data

It may seem as if the problem is just one of insufficiently varied data. After all, research into the nature of some kind always needs a heterogeneous data base. If we lack sufficiently varied data points, we are in danger of picking out some narrower or wider category than the target kind itself. So for example, if all our examples of pain were stabbing pains, we might be led to posit an overly precise material nature for pain, by fixing on the material basis of stabbing pains, rather than of pain as such. Conversely, if none of our examples of pain absence were heat sensations, we may be led to an overly permissive material nature for pain, by fixing on some material property common to both pain and heat sensations. To get things right, our sample must thus include sufficiently varied examples of pains and non-pains. (For a careful discussion of how to maximize such variation in research on humans, see Frith et al., 1999.)

Isn't just this our present difficulty? You might think that if only we had appropriately varied examples of pains and non-pains, then we could decide between P and S as the material nature of pain easily enough. In normal humans, P and S always go hand in hand, so normal humans won't supply the requisite data points. But if we could find some beings who have P without S, or vice versa, and could see whether they still have pains, then that would do the trick.

In fact there is only a possibility of dissociation in one direction here, since the carbon-based details P will determine the structural property S. Still, we can well have S without P, as in the silicon doppelganger. And this doppelganger itself might then seem to offer the requisite data point, in principle at least. We simply need to check whether the silicon doppelganger with S but not P is in pain. If it is in pain, then this will eliminate the carbon-based P as unnecessary for pain. And if it isn't in pain, then that will eliminate the structural S as insufficient for pain.

But this doesn't work. There is a deeper difficulty, which will remain even after we incorporate a silicon doppelganger into our research. The trouble is that our canonical method for telling whether or not we have a case of pain is to rely on subjects' reports. In the case at hand, we will thus ask the doppelganger whether or not it consciously feels pain when it is pricked with a pin. However, we know beforehand what the doppelganger will say, whether or not it actually feels any pain. Since it is structurally organized just like us, it will behave

observably just as we do, and in particular will utter the words 'I am in pain' when it is pricked with a pin. But this tells us nothing at all, since we already know it will utter these words both in the case where it lacks real pain and in the case where it has it.

Clearly, it would beg the question at issue to take the doppelganger's utterance at face value. Our question is whether structural S or physical P constitutes the material nature of pain. If structural S did constitute pain, then the doppelganger would feel pain, and its words 'I am in pain' would report as much. On the other hand, if physical P constituted pain, then the doppelganger wouldn't feel pain, and its words would have some other significance. So, in order to interpret the doppelganger's words, we need to know whether S or P constitutes pain, which was the question we started off trying to answer.

VI What's So Special About Humans?

At this point an obvious question may occur to some readers. If we can't take the doppelganger's reports at face value, then why isn't there a similar problem with the phenomenal reports of other humans? As I have described the methodology of consciousness research, we are looking for a property which is present whenever humans say 'I am consciously experiencing phenomenal property \emptyset ' and which is absent whenever they deny this. This methodology thus takes it for granted that the phenomenal reports of other humans provide accurate information about \emptyset . But why is there an asymmetry between humans and the doppelganger here? Shouldn't we be suspicious of the reports of other humans too?

My answer is that when we are dealing with other humans, rather than silicon doppelgangers, we have good independent reason to suppose that they share our experiences, and thus that they refer to those experiences when they form phenomenal concepts by 'quoting' their own experiences.

This might seem rather quick. What about the familiar challenge of the inverted spectrum? Isn't it possible that, when you look at lush grass, you have the experience that I have when I look at ripe tomatoes?

Well, I concede that this is a priori possible, even given our similar behavioural responses to colours. But there is also a posteriori reason to think that such inverted spectra are not actual. I have in mind here the general empirical arguments in favour of materialism mentioned at the beginning of this paper: given the (a posteriori) causal completeness of physics, we have good empirical reason to reject the metaphysical dangles of dualism in favour of the relative simplicity of materialism; so conscious properties must be determined by physical properties; which means we can be sure that any beings who share relevant physical properties will share conscious properties. Since we also have empirical reason to suppose that normal humans are alike in relevant physical details when they look at ripe tomatoes, we can conclude that ripe tomatoes occasion the same experience in all normal humans.

In support of this line of argument, note how we would not rely on the phenomenal reports of human beings whom we knew not to be physically similar in relevant respects. Thus consider a brain-damaged human who has had its cerebral pain mechanisms replaced by some structurally equivalent but silicon-based prosthesis. Even though this is unequivocally a human being, it will count with the silicon doppelganger as far as consciousness research goes: it is no good asking this person whether it feels a pain when it is pricked, in the hope of discovering whether S (which it has) or P (which it lacks) is the material basis of pain. For its structural equivalence with normal humans ensures that that it will respond 'Yes', whether or not it feels pain. (Similarly, it is no good engineering somebody in whom my 'red' brain state realizes the detailed 'green' structure, and then asking it whether it sees something green when it looks at grass, in the hope of discovering whether colour experience goes with structure or realizer. For again this person will say 'Yes', whether or not it sees something green.)

Indeed a similar point would apply if it were to turn out that normal healthy humans varied in the material underpinnings of their conscious states. Suppose, just for the sake of the argument, that we discovered that the pain mechanisms of men and women were realized by somewhat different neuronal arrangements, even though these mechanisms were computationally identical, with both sexes agreeing entirely in their subjective reports about pain. Then there would be some computational structure that co-varied with pain in both men and women, while the neuronal realizers of this structure differed between men and women. In this situation, women could take the subjective pain reports of other women at face value, and be confident that other women meant the same as them when they said 'I am in pain', on the grounds that women shared all relevant material properties. But women could not take the same attitude to male pain reports, since they could not rule out the possibility that pain experience depended on neuronal underpinning rather than computational structure, and that men were therefore referring to something different from female pain with their phenomenal concept of 'pain'.

VII Broad Alternatives

So far I have argued that scientific research will be impotent to decide between strictly physical and structural properties as the referents of phenomenal concepts. Let me now show that a similar problem arises, not only with 'vertical' alternatives among narrow material properties, but equally with 'horizontal' alternatives that emerge once we entertain the possibility that conscious properties may be broad.

The point is best introduced by considering views of consciousness which are not just materialist but also representationalist—that is, views which seek to equate phenomenal properties with representational properties, where representation is somehow explained in material terms. Now, most current materialist accounts of representation make representation a broad matter: to represent that there is water in your glass, or that an elephant is approaching, or that your wife is leaving, it is not enough that your brain be appropriately internally arranged—you also need to have appropriate causal or historical relationships to water, or elephants, or your wife. A physical duplicate on a distant planet would not be able to represent these things.

At first pass, representational accounts of phenomenal properties thus appear to have the surprising consequence that your phenomenal states, along with your other representational states, depend on your causal or historical relations to your environment, and not just on current matters inside your skin. Some philosophers of consciousness are prepared to swallow this (Dretske, 1995, Tye, 2002). However, there is also the option of factoring out a narrow component of representational properties—which depends only on matters inside the skins of water/elephant/wife representers—and equating phenomenal properties with these narrow representational properties instead.

So any phenomenal property \emptyset will once more present us with a choice: should \emptyset be equated with some broad representational property B, or with its narrow distillation N? Well, we could decide this empirically if we had an effective test case. We need a being with N but not B, or vice versa, and a way of telling whether it has \emptyset .

Again there is only a possibility of dissociation in one direction here, since broad representational properties will include their narrow components. However, we can imagine a being who is internally just like you, but lacks your causal or historical environment (it lives on a distant planet, say, or was constituted by chance in a swamp). So all we need to know is whether this being feels \emptyset when you do. But now we face the same old problem. The way to find out whether it feels \emptyset is to ask it. But we know beforehand what it will say. Since it is just like you in narrow respects, it will certainly utter the sounds 'I am feeling \emptyset '. Yet it would utter these sounds even if phenomenal properties were broad, and it did not feel \emptyset at all. Once more, empirical research cannot decide between the alternative material referents for phenomenal concepts.

VIII. The Issue In General Terms

It may be helpful at this point to move away from details, and lay out the dialectical situation in general terms. I began with abstract considerations resting on the causal completeness of physics: these implied that any given conscious property \emptyset must be identical to some material property or other. The identification of some specific material property was then turned over to empirical investigations. Ideally, we would like the abstract considerations and the empirical investigations to complement each other, with the empirical investigations identifying some unique material property M that is perfectly correlated with phenomenal property \emptyset in humans, and the abstract considerations then allowing us to conclude that M and \emptyset aren't just correlated, but identical.

The trouble, however, is that empirical investigations seem unable to play their part. They can't identify a unique material property, since there are too many material properties which are perfectly correlated in humans with any given phenomenal \emptyset —structural properties as well as strictly physical ones, broad as well as narrow. So empirical investigations seem unable to pinpoint the material referents of phenomenal concepts.

Some readers may feel that the way forward is to beef up the abstract considerations. Perhaps we are asking too much of the empirical investigations. Rather we should first seek to decide, on general abstract grounds, whether we are looking for a structural or a strictly physical property, or a broad or a narrow one. Then, once the target has been narrowed down in this way, there need be no barrier to empirical research locating unique material referents for phenomenal concepts.

So perhaps we could decide against structural properties, and in favour of strictly physical ones, on the grounds that only strictly physical occurrences have the kind of strong causal efficacy appropriate to conscious causes (cf. Kim, 1998). Again, perhaps we could decide against broad properties, and in favour of narrow ones, on the grounds that only narrowness is consistent with the authority which subjects have over their conscious states (cf. Block, 1990).

But to my mind this is asking rather a lot of the abstract arguments. While the completeness of physics does argue strongly against metaphysically independent dualistic properties, it does nothing to rule out metaphysically dependent structural properties, and it is by no means clear that these structural properties are lacking in the kind of causal strength that is possessed by conscious properties (Papineau, 2002, section 1.7). Again, it is not obvious that first-person authority about conscious states is incompatible with their external

constitution; an analogous compatibility has been widely defended by externalists about propositional attitudes, and externalists about consciousness have appealed to similar arguments (Tye, 2000).

In any case, this is not the place to pursue these issues. As I said, I am pessimistic. So I shall assume, for the sake of the subsequent argument, that general abstract considerations are not sufficient to decide whether conscious properties are structural rather than physical, or broad rather than narrow, or between other similar alternatives. My main point so far has been that, if I am right about the impotence of general abstract arguments, then it is no use turning to empirical science, for that won't help us either.

XI Vagueness

If abstract analysis combined with empirical investigation cannot identify unique material referents for phenomenal concepts, does this mean that we must remain permanently ignorant of the basic material facts about consciousness? Is there some epistemological barrier that stops us finding out whether the silicon doppelganger is actually in pain?

Not at all. I do not take the above arguments to show that there are real truths about consciousness which lie beyond our epistemological access. Rather, the underlying problem is that our phenomenal concepts are vague. Because of this, there is simply no fact of the matter as to whether the doppelganger is in pain. Our phenomenal concept of pain is not focused enough to decide the issue. Nothing in the semantic constitution of the concept decides whether it refers to the physical mechanism P which realizes pain in humans, or to the structural S which we humans share with the doppelganger. So not even God, who knows everything, would be able to tell whether or not the doppelganger with S is in pain, for roughly the same reasons that God would be unable to tell whether I am bald or not.

Actually, an analogy with baldness is less than perfect, since baldness is a matter of degree, vague because there is no sharp cut-off point. Phenomenal concepts, by contrast, are vague because it is indeterminate what kind of property they refer to, rather than where to divide some continuum. A better analogy is provided by the traditional Eskimo word for whale oil (Block, forthcoming). At some point in Eskimo history, a petroleum product which looked and functioned just like natural whale oil was introduced as a substitute. Did the original Eskimo word (let's pretend it was 'whale oil') apply to this new substance or not? If 'whale oil' referred to a biologically or chemically identified type, it did not; but if 'whale oil' referred to anything with the requisite appearance and use, then it did. But of course there is likely to be no answer here. Nothing in previous Eskimo linguistic practice or usage need have determined which way 'whale oil' should be read when this question arose. And then there would be no fact of the matter as to whether the new stuff is whale oil. Similarly, I say, with pain. There is no fact of the matter as to whether the doppelganger is in pain, because nothing in previous usage fixes whether it should apply to physical P or to structural S when the question is raised.

Now, this line on consciousness would be uncomfortably ad hoc if my only reason for saying that phenomenal concepts are vague were that we can't find answers to the questions they pose. It is a mistake to infer an indefiniteness of questions directly from an unavailability of answers. This would require the mistaken verificationist doctrine that our concepts cannot lay claim to matters beyond our epistemological access. However, I think that there are good independent reasons for taking phenomenal concepts to be vague.

I take phenomenal concepts to be basic unstructured terms, whose referential powers depend on the information they (are designed) to causally track (Papineau, 2002, chapter 4). At bottom, I take this to be a matter of tracking the experiences of human beings. But experiences are material, if they are anything, and, as we have seen, there are various different material properties, physical and structural, broad and narrow, which will serve to make the same classifications among normal human beings. So there is no reason to suppose that there is anything in the semantic constitution of phenomenal concepts to focus them precisely on one rather than another of these humanly coextensive material properties—just as there is no reason to suppose that anything focuses the Eskimo term for 'whale oil' on one possible referent rather than another.

No doubt some readers will feel that this analogy with 'whale oil' is misplaced. In the whale oil case, it seems clear that nothing is at issue but a matter of conceptual refinement. If we read 'whale oil' as referring to a chemical substance, then the new stuff doesn't count as 'whale oil'—but if we go the other way, it does. The pain case, by contrast, doesn't seem like a matter of conceptual decision at all. True, if we were to read 'pain' as referring to the structural S, then the doppelganger would be in the extension of 'pain', and if we were to read 'pain' as referring to the physical P, it wouldn't be. But in this case we feel there remains a further question. We want to know whether the doppelganger is in pain, not how to apply some term, and we feel that any decision as to whether 'pain' refers to S or P should respect this concern, with 'pain' referring to whichever material property in fact fixes real pain.

Still, I want to stand by the analogy with the whale oil case. In my view, the pain case seems different only because we find it so natural to think about conscious experiences dualistically. We unreflectively assume that pain is metaphysically distinct from any material properties like S or P. And once we do this, then of course

we take there to be a fact of the matter as to whether the doppelganger has this distinct property or not. But this is all a mistake. There isn't any distinct pain property, different from physical and structural material properties, or broad and narrow ones, any more than there is a distinct whale oil property, different from chemical constitution, biological origin, appearance and use. So the person who insists on asking whether the doppelganger is in pain is like an Eskimo who says 'Yes, I can see that this new manufactured stuff is like the old stuff in appearance and use, while different in chemical constitution and biological origin. But that doesn't yet answer my question, which was whether it is like the old stuff in being made of whale oil.' This person would be assuming that being made of whale oil is some extra property, over and above constitution, use and so on. Similarly, the persistent desire to know whether the doppelganger is really in pain stems from the assumption that pain is some extra property, over and above physical and structural ones.

Why is it so compelling to suppose that conscious experiences involve extra properties, over and above material properties, when this seems absurd for whale oil? This is a large question, which I shall not even attempt to answer here. My Thinking About Consciousness (2002) is a sustained attempt to locate the source of this near-inescapable 'illusion of distinctness'. For the present, let me simply state that I take this illusion to stem from the special structure of phenomenal concepts, and in particular from the way they 'quote' the experiences they refer to: this quotational feature makes us feel that other, material ways of referring to experiences 'leave out' the experiences themselves, and thus that the phenomenal concepts must refer to something extra non-material property.

XII Relation to Earlier Debates

Philosophical debates about alternative material referents for mental terms are not new. The question of whether terms like 'pain' refer to higher-order role properties or to realizer properties has been a staple of analytic philosophy of mind for some decades. (See the essays in Block (ed.), 1980.) Similar issues about broad and narrow properties are also familiar. However, these debates are normally taken to be about the referents of 'psychological' concepts defined in terms of theoretical roles. In this paper I have focused instead on the referents of phenomenal concepts. I think it is helpful to recognize that psychological and phenomenal concepts raise different issues.

If our only concern were the referents of 'psychological' role concepts, then it would be quite mysterious why 'pain' strikes us so differently from 'whale oil'. The psychological concept of pain is defined in terms of canonical causes like bodily damage and effects like avoidance behaviour. Now, does the term refer to any being displaying this structure of causes and effects, or only to those who share the physical property that mediates this structure in humans? So far this looks like a simple question for decision, quite analogous to the 'whale oil' case—after all, concepts defined in terms of theoretical roles can be taken either way, so what more is needed but a decision about how to use the term 'pain'?

To make sense of the importance attached to these debates, we need to read them as focused on the referents of phenomenal concepts. This is because it is specifically phenomenal concepts that give rise to the illusion of extra, non-material referents, in the way indicated at the end of the last section. And it is this illusion that makes us feel that there must be something more to our questions about reference than opportunities for terminological decision. It is specifically because we think that phenomenal pain involves something metaphysically independent of physical nature that we feel that the doppelganger issue cannot simply be settled by terminological fiat alone.

XII What It Is Like

Does it really make sense to hold that it is vague whether the doppelganger is in pain? How can this be? Either the doppelganger feels a certain way or it does not. It can't be vague how it feels.

But I do not want to argue that there is something less than definite in the doppelganger's experience itself. The doppelganger definitely feels as it does, however that is. My thesis is only that it is indeterminate whether it is pain, where I take this indeterminacy to arise from a vagueness in our human term 'pain'. The indeterminacy doesn't lie in the doppelganger's experience itself, but in whether that experience is similar enough to cases of human pain to fall under our term 'pain', a term whose content derives from exemplars provided by human pains.

At this point some readers may want to ask about the general phenomenal concept of consciousness-as-such, as opposed to specific phenomenal concepts like pain or seeing something red. Maybe it makes sense to hold that it is vague whether doppelgangers count as in 'pain', on the grounds that it is unclear how similar their experiences are to human pain. But surely I don't want to say whether it is vague whether they are conscious at all. Mustn't this at least be definite, however it is with their more specific modes of experience?

But I do want to say that even phenomenal consciousness-as-such is vague. So far in this paper I have focused on specific phenomenal concepts like pain. But everything I have said about these applies equally to the general phenomenal concept conscious-as-such.

As I observed at the beginning of this paper, there are no a priori connections between the phenomenal concept conscious-as-such and any material concepts. So it will be matter of empirical research to identify the material referent of the phenomenal concept conscious-as-such. However, when we look for some material property common to all those cases where humans report they are phenomenally conscious, and absent whenever they deny this, we will inevitably find a variety of material candidates which fit perfectly with this data base. Apart from any strictly physical property, there will also be various structural properties. Again, there will be broad properties as well as narrow ones. In addition—though this is not the place to go into details—there will be candidates involving Higher-Order Thoughts as well as candidates involving only lower-order attentional properties (Papineau, 2002, section 7.11 to 7.13).

Just as before, I do not take this to show that we are trapped in ignorance about the true material nature of consciousness-as-such. Rather, the fault lies with our phenomenal concept of this property, which is vague in the crucial dimensions. We somehow construct this concept by taking exemplars of human conscious states, and use the resulting concept to classify humans depending on whether or not they are in some such conscious state or not. There is no reason to suppose the concept so constituted is precisely enough focused to discriminate between the physical and structural properties shared by conscious humans, or between the broad and narrow properties, or between the Higher-Order and attentional properties.

Still, readers might wish to press me, how can it be vague whether a silicon doppelgangers is conscious at all? (Surely a light is on, or it is not.) But, as before, my thought is not that it is vague how it is for the doppelganger itself. Rather, the vagueness lies in our concept, which fails to decide whether the doppelganger's mode of being is sufficiently like our own to count as 'conscious' or not.

But, you may wish to persist, isn't it either like something to be the doppelganger, or not? How can that be vague? (Surely a light is on, or it is not.) But I am sceptical as to whether the phrase 'like something to be' will bear the weight of this insistence. There are many modes of being apart from our own, including those of cats, fish, slugs, plants, stones, possible silicon humanoids, and no doubt many others. I see no reason to suppose that the phrase 'like something to be' will draw a sharp line across this range, with some beings coming out as clearly conscious, and the rest not. The various cases will all display some material similarities with human consciousness, and some differences. Given this, there will be nothing to determine whether they fall under our term 'conscious' or not.

XIII Conclusion

Over the past decade or so, there has been a growing wave of enthusiasm for scientific research into consciousness. After a long period when consciousness was regarded as beyond the scientific pale, many scientists now aim to penetrate the mysteries of consciousness and identify the material mechanisms that underlie conscious experience. If the argument of this paper is sound, however, their hopes will at best be partially fulfilled.

The most we can expect from scientific research is the identification of a range of material properties that correlate in human beings with pain, say, or colours, or indeed being conscious at all. However, this won't pinpoint the material essence of any such conscious state, for there will always be a plurality of such human material correlates for any conscious property. And in the face of this plurality science is impotent, for the methodology of consciousness research offers no handle by which to separate material properties that are perfectly correlated in normal humans.

So scientific research into consciousness is fated to disappoint its more extreme enthusiasts. This does not mean, however, that we would do better to turn to some alternative mode of investigation. It is not as if conscious properties have true material essences, yet science is unable to discover them. Rather the whole idea of identifying such essences is a chimera, fostered by the impression that our phenomenal concepts of conscious states are more precise than they are.

References

Block, N (ed.). 1980. Readings in the Philosophy of Psychology, vol. 1. London: Methuen.

Block, N. 1990. 'Inverted Earth', in Philosophical Perspectives 4, ed. J. Tomberlin. Northridge, Calif: Ridgeview.

Chalmers, D. 1996. The Conscious Mind. Oxford: Oxford University Press.

Dennett, D. 1978. 'Why You Can't Make a Computer That Feels Pain', in his Brainstorms. Cambridge, Mass: Bradford Books.

Dretske, F. 1995. Naturalizing the Mind. Cambridge, Mass: MIT Press.

Frith, C., Perry, R., and Lumer, E. 1999. 'The Neural Correlates of Cognitive Experience: An Experimental Framework'. Trends in Cognitive Neuroscience, 3.

Goodale, M. and Milner, A. 1992. 'Separate Visual Pathways for Perception and Action'. Trends in Neurosciences, 15.

Jackson, F. 1986. 'What Mary Didn't Know'. Journal of Philosophy, 83.

Kim, J. 1998. Mind in a Physical World. Cambridge, Mass: MIT Press.

Kripke, S. 1980. Naming and Necessity. Cambridge, Mass: Harvard University Press.

Lycan, W. 1987. Consciousness. Cambridge, Mass: MIT Press.

Papineau, D. 2002. Thinking about Consciousness. Oxford: Oxford University Press.

Tye, M. 2000. 'Inverted Earth Meets Swampman', in his Consciousness, Color and Content. Cambridge, Mass: MIT Press.

Weiskrantz, L. 1986. Blindsight. Oxford: Oxford University Press.