Chapter 2

# Explanatory gaps and dualist intuitions

David Papineau

## 2.1 Introduction

I agree with nearly everything Martin Davies says. He has written an elegant and highly informative analysis of recent philosophical debates about the mind–brain relation. I particularly enjoyed Davies' discussion of B.A. Farrell, his precursor in the Oxford Wilde Readership (now Professorship) in Mental Philosophy. It is intriguing to see how closely Farrell anticipated many of the moves made by more recent 'type-A' physicalists who seek to show that, upon analysis, claims about conscious states turn out to be nothing more than complex third-personal claims about internal and external behaviour. Davies is also exemplary in his even-handed treatment of those contemporary 'type-B' physicalists who have turned away from the neo-logical-behaviourism of Farrell and his ilk. Davies explains how type-B physicalists recognize distinctive subjective 'phenomenal concepts' for thinking about conscious states and so deny that phenomenal claims can be deduced *a priori* from behavioural or other third-personal claims. However, type-B physicalists do not accept that these subjective phenomenal concepts refer to any distinct non-material reality. In their view, phenomenal concepts and third-personal scientific concepts are simply another example of the familiar circumstance where we have two different ways of referring to a single reality.

Since I am persuaded by pretty much all Davies says about these matters, I shall not comment substantially on the dialectical points he covers. Instead I want to raise two rather wider issues. The first is the set of ideas associated with the phrase 'the explanatory gap'. Davies specifies that he is using this phrase in a specific technical sense. But the phrase has further connotations, and this can lead to a distorted appreciation of the philosophical issues. The second issue is the methodological implications of the philosophical debate. I shall argue that the philosophical issues addressed by Davies suggest that there are unexpected limitations on what empirical brain research can achieve.

In what follows, sections 2–5 will be devoted to 'the explanatory gap', while the rest of my remarks will address the methodological issue.

## 2.2 **Terminology**

Along with many other contemporary philosophers of consciousness, Davies takes the phrase 'the explanatory gap' to refer to the absence of any *a priori* route from the physical facts to the conscious facts. Davies agrees that, however detailed a physical description of some system we are given, it will never follow *a priori* that the system must have certain conscious experiences. Even if it is specified that some being is physically exactly like one of us humans, it will remain conceivable that this being is a 'zombie' with no experiences at all. As Davies uses the phrase, 'the explanatory gap' thus refers to the lack of any conceptual tie between physical descriptions and claims about consciousness.

Given this terminology, type-A and type-B physicalists can be distinguished by their differing attitudes to this 'explanatory gap'. Type-A physicalists will insist that at bottom there is no 'explanatory gap' and that, despite appearances to the contrary, claims about consciousness can after all be shown to follow *a priori* from the physical facts. Type-B physicalists, by contrast, accept the 'explanatory gap', but maintain that no ontological conclusions about the immateriality of consciousness follow.

At one point (his section 2.4) Davies considers Michael Tye's objection to using the phrase 'explanatory gap' for nothing more than the lack of a conceptual tie between brain and mind. Tye complains that, if there is no conceptual connection between the physical and mental realms, then this isn't the kind of lacuna that that can be filled by future empirical discoveries, and that therefore it is odd to call it an 'explanatory gap', as if it were something that will go away once we get hold of the right empirical information.

Davies immediately concedes to Tye that type-B physicalists who recognize distinctive phenomenal concepts of conscious states will hold that the conceptual lacuna is here to stay. Davies also agrees with Tye that this isn't necessarily fatal to the type-B position. But he opts to keep the term 'explanatory gap' for the lack of a conceptual tie between the physical and mental realms.

Davies says that this is just a terminological matter, and of course he is entitled to specify how he is to be understood. But even so I think there is a real danger hiding behind the usage he endorses. The trouble is that there is something else that makes people posit an 'explanatory gap' between brain and mind, quite apart from the lack of conceptual brain–mind ties. Moreover, this something else threatens physicalism far more directly than any difficulties raised by the conceptual lacuna. Given this, Davies' preferred usage can easily create the impression that the conceptual issue is more important than it is.

## 2.3 **The intuition of distinctness**

I think that the real reason most people feel that there is an 'explanatory gap' between the physical and conscious realms is simply that they find physicalism incredible. Physicalism about consciousness is a strong claim. It isn't the relatively anodyne claim that conscious states are closely connected with what is going on in the brain, in the way that smoke is connected with fire. Rather physicalism says that conscious states are brain processes, in the way that water is $H_2O$. Smoke is caused by fire. But water isn't caused by $H_2O$—it is $H_2O$. Similarly, according to physicalism, pain isn't caused by a brain process—it is a brain process.

This is a very hard thing to believe. The more intuitive thought is surely that conscious states are extra to brain processes, even if they are closely correlated with them. Elsewhere I have called this natural dualist thought the 'intuition of distinctness' (Papineau 2002.) I would say that it strikes most people as obvious that the conscious mind is something more than the brain and that physicalism is therefore false.

Indeed I would say that there is a sense in which even professed philosophical physicalists, including myself, cannot fully free themselves from this intuition of distinctness. Of course, we deny dualism in our writings, and take the theoretical arguments against it to be compelling. But when we aren't concentrating, we slip back into thinking of conscious feelings as something extra to the brain, in the way that smoke is extra to fire, and not as one and the same as the brain, in the way that water is one and the same as $H_2O$.

I take this to be the real lesson of Saul Kripke's celebrated argument at the end of *Naming and Necessity* (1980). Kripke points out that even 'identity theorists' will admit that mind–brain identity claims appear contingent: even professed physicalists admit that it seems possible that pain could have failed to accompany the brain processes with which it is inevitably found in this world. Kripke then argues, and I think he is quite right, that this appearance can't be explained consistently with the professed physicalism. If you really believe that pain is one and the same as some brain process, then how can it so much as seem to you that pain and that selfsame brain process might have come apart in some other possible world? How can something come apart from itself? The moral of Kripke's argument, as I read it, is thus that even professed physicalists don't always fully believe their physicalism. They slip into thinking of pain as something that accompanies the relevant brain process, rather than as identical to it, and that's why it appears to them that the two things may come apart. (It should be said that this isn't the standard reading of Kripke. For more on this see my 2007.)

There is also some rather more immediate evidence that professed physicalists don't always believe their physicalism, even before we bring in Kripke's sophisticated modal considerations. The very language used by physicalists often gives the game away. Consider the phrases normally used to raise questions about the relation between the brain and the conscious mind. What are the 'neural correlates' of consciousness? Which brain states 'accompany' conscious states? Again, how do brain processes 'generate' conscious states, or 'give rise to' them? Such questions are commonly posed in discussions of consciousness by many people who would adamantly insist that they reject dualism. But their phraseology shows that they are not consistent in this denial. If they really thought that conscious states are one and the same as brain states, they wouldn't say that the one 'generates' or 'gives rise to' the other, nor that it 'accompanies' or 'is correlated with' it. $H_2O$ doesn't 'generate' water, nor does it 'give rise to' or 'accompany' or 'become correlated with' it. $H_2O$ simply is water. To speak of brain states as 'generating' consciousness, and so on, only makes sense if you are implicitly thinking of the consciousness as ontologically additional to the brain states.

The obvious question to ask at this point is why physicalism should strike us all as intuitively false, even if it is true. This is a very interesting issue, on which a number of writers have expressed views (Papineau 1993, 2002, 2006; Melnyk 2003; Bloom 2004).

One possibility is that the intuition is forced on us by some deep feature of our cognitive system. Alternatively, it may be a relatively superficial phenomenon, stemming from the theoretical presuppositions of contemporary culture, and so likely to disappear if those presuppositions change. However, this is not the place to pursue this topic. Let me content myself by pointing out that, even if the intuition of distinctness is here to stay, it is something that theoretical physicalists can happily live with.

After all, there are plenty of other cases where our best theoretical beliefs conflict with deep-seated and unmalleable intuitions. Consider the familiar Müller–Lyer illusion. At a theoretical level, we know that the lines are the same length. But at a more intuitive level of judgement they strike us as of different lengths. Nor is this kind of set-up restricted to cases involving perceptual illusion. At a theoretical level, I am entirely convinced that there is no moving present and nothing is left out by a 'block universe' description of reality which simply specifies the dates at which everything occurs. But at an intuitive level I can't stop myself thinking that I am moving through time. At a theoretical level, I am persuaded that reality splits into independent branches whenever a quantum chance is actualized. But at an intuitive level I can't shake off the belief that there will be a fact of the matter about whether the Geiger counter will click in the next two seconds or not.

I take it that in all these cases it is clear that, whatever the source of the contrary intuitions, they do not constitute serious obstacles to the theoretical beliefs. Similarly, I say, with consciousness. We may be unable to shake off the intuitive feeling that mind and brain are distinct. But that is no substantial reason to doubt the theoretical reasons for identifying them.

## 2.4 **The feeling of a gap**

In my view, the intuition of dualist distinctness is the real reason for the widespread feeling of an 'explanatory gap'. This feeling is nothing to do with the lack of conceptual ties between physical and conscious claims. When people feel that a 'why' question remains even after the 'what' question of consciousness has been answered, as Davies puts it in his introductory remarks, this isn't because they are worried that they can't deduce the conscious facts *a priori* from the physical facts. It's simply because they are assuming that conscious states are ontologically distinct from brain states. And this assumption of course gives rise to an obvious and urgent explanatory question—why do certain physical processes have the mysterious power of extruding a special mind-stuff?

Still, this question is generated directly by the intuition of dualism, and not by the relatively esoteric circumstance that conscious facts can't be deduced *a priori* from the physical facts. The best way to see this is to note that there are plenty of other cases where such *a priori* deductions are not available, yet we don't have any feeling of a worrying explanatory gap. Davies himself offers one kind of example: identities like that 'location L is here' (where location L is somehow specified objectively). Clearly there isn't any question of *a priori* deducing the here-facts (there is water here, say) from the objective spatial facts (such as that there is water at location L) however many objective spatial facts we are given. Yet we feel no explanatory gap in this case (we don't ask why is location L here?). Or, again, consider identities framed with the help of proper names, such as that Kripke is the greatest living philosopher. I take it that there is no question of deducing this claim *a priori* from the physical facts, or any other set of facts stated without using the name 'Kripke'. Yet surely we don't on that account feel that we lack some further explanation of why the greatest living philosopher is Kripke.

I infer that the real reason people want to say there is an 'explanatory gap' in the mind–brain case is that they can't help thinking in dualist terms—they want to understand why certain physical processes extrude the extra mind-stuff. If the only problem were that they can't deduce the conscious facts *a priori* from the physical facts, they wouldn't experience any worrying 'explanatory gap'. They would be no more puzzled about the mind–brain case than they are about why location L is here, or why Kripke is the greatest living philosopher.

## 2.5 **Dangers of confusion**

Davies makes it clear that he is using the phrase 'the explanatory gap' to refer specifically to the lack of any conceptual tie between brain and conscious mind and not to any dualist intuition. And, as I said, it is up to him to stipulate how he is going to use words. Still, I hope it is now clear why I think there is some danger that this usage may mislead unwary readers.

After all, talk of 'the explanatory gap' will inevitably be understood by most people as expressing the dualist intuition that the conscious mind is separate from the brain. Yet at the same time Davies, along with many others, specifies that this phrase is to be understood as referring to the conceptual separation between brain and conscious mind. This can scarcely fail to create the impression that this lack of a conceptual tie provides prima facie support for dualism. To my mind, of course, it does no such thing. As I have just explained, there are plenty of other cases apart from the mind–brain relation where a conceivability gap is manifestly compatible with identity. So from my perspective it is little more than a coincidence that the mind–brain case is associated with both a conceivability gap and an intuition of ontological distinctness. The two things really have nothing to do with each other. Still, it can be hard to see this, if we refer to the conceivability gap using a phrase that is standardly understood as expressing the dualist intuition.

In effect, my worry is that Davies' preferred terminology gives the Jackson–Chalmers line of argument an unwarranted rhetorical boost. Jackson and Chalmers hold that there is a sound argument from the lack of a conceptual tie to dualism (e.g. Chalmers and Jackson 2001). This is a serious and influential thesis, and there is no doubt that it deserves the careful yet critical attention that Davies gives it. But I wonder whether it would carry quite the same initial plausibility, were it not for the widespread practice of describing the premise with a phrase that is naturally understood as expressing the conclusion.

I can make a related point about explanation itself. In stipulating that the 'explanatory gap' signifies the absence of conceptual ties between physical and conscious realms, Davies and others who adopt this usage strongly suggest that something will be left unexplained if such conceptual links cannot be forged. The very phraseology implies that a type-B physicalism that tries to live without such ties will inevitably leave us with unfinished explanatory business. This view is certainly supported by Davies' opening remarks, where he urges that even an answer to the question of 'what' physical states correlate with consciousness will fail to resolve the 'intractable' question of 'why' they do so.

But this casts an unwarranted aspersion on type-B physicalism. From the perspective urged here, it is a mistake to think it leaves anything unexplained. Of course, as I said earlier, if you slip into hearing the 'explanatory gap' as

expressing an intuition of dualist distinctness, then of course you will think it raises an urgent explanatory demand—the need to explain why some special physical processes should extrude a distinctive mind-stuff. But any clear-thinking type-B physicalist will resist this dualist intuition, and then it is by no means clear that there is anything left to explain.

We have already seen reason to doubt that a conceptual lacuna *per se* generates explanatory demands. There are no conceptual ties between objective spatial facts and egocentrically identified ones, yet we don't feel we need to explain 'why location L is here'. Again, there is no conceptual tie between descriptive and proper name facts, yet we don't feel called to explain why 'the greatest living philosopher is Kripke'. I say the same about the mind–brain case. There may be no conceptual route from brain descriptions to claims about consciousness, but that doesn't mean non-dualists need to explain why given brain states are the conscious states they are.

Claims about identity or constitution don't need explanation, even when they aren't conceptually guaranteed. If two things accompany each other, we can rightly ask why this is so. But if something is the same as something else, it makes no sense to continue asking for an explanation. If you think that Samuel Clemens and Mark Twain are two different people, you will rightly be puzzled why they are always in the same place at the same time. But once you realize the truth, you won't go on asking why Samuels Clemens is Mark Twain. Similarly, if you think that pains accompany C-fibre stimulation, you might well wonder why this is so. But once you accept that they are one and the same state, there is no remaining need to explain why that state is itself.

From my point of view, the so-called 'explanatory gap' does not indicate a failing in type-B physicalism, so much as a failing in many professed type-B physicalists. The trouble isn't that type-B physicalism is somehow incomplete, but rather that many type-B physicalists fail to believe it wholeheartedly. They slide back into dualist thinking, and then find themselves hankering for explanations.

So the right reaction to talk of the 'explanatory gap' isn't that type-B physicalism somehow needs fixing or supplementing. It is just fine as it is. The difficulty is that we don't commit ourselves properly. If only we fully believed our physicalism, our explanatory hankerings would die away.

## 2.6 **Methodological limitations**

I turn now to the methodological implications of type-B physicalism. As we have seen, type-B physicalists explain the relation between mind and brain by positing special 'phenomenal concepts'. These concepts have no *a priori* connections with any scientific third-personal concepts. Still, in reality they

refer to nothing except brain processes that can also be referred to using scientific concepts. So type-B physicalists combine ontological monism with conceptual dualism. There is just one physical reality, but we can think about the conscious part of that reality in two ways, either with third-person scientific concepts or with subjective phenomenal concepts.

However, once we bring in these special phenomenal concepts, then it turns out that they allow us to pose certain questions about consciousness that are very difficult to answer. Indeed I don't think that they can be answered at all. However, I don't think that this points to any real epistemological limitation facing consciousness research. The trouble is rather that phenomenal concepts are not precise enough to frame all the questions that we intuitively expect them to frame.

## 2.7 **Subjects' reports**

We can bring out the problem by thinking about the methodology of consciousness research. How do we find out which brain processes constitute pain, or consciously seeing something, or indeed being conscious at all? The obvious technique is to appeal to subjects' reports. We take human subjects in given conditions, ask them whether or not they were in the relevant conscious state, and seek to find some brain state that is present whenever they say 'yes' and absent whenever they say 'no'.

I shall have more to say about the methodological significance of subjects' reports below. But we can bring out the way in which consciousness research normally hinges on such reports by considering subjects who are not capable of making reports, like vervet monkeys, say. You can submit vervet monkeys to various experimental conditions, you can get them to perform various tasks, and you can check on what is going on in their brains at such times. But since the monkeys can't tell you what they consciously experience, 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 explicitly tell us whether or not they are feeling a pain, seeing something red, or experiencing something rather than nothing, and this offers a crucial handle by which to identify the material nature of phenomenal properties.

## 2.8 **Too many candidates**

Now, there is no doubt that research based on such reports by human subjects 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 pretheoretically have expected to be conscious: these include the visual guidance of object-oriented 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).

However, such research can only take us so far. It can narrow down the candidates for the physical nature of given conscious states, but it can never identify the relevant physical state uniquely. For example, we can be sure that more is required for conscious vision than the processes guiding object-oriented hand movements, and that less is required for conscious pain than the processes present in unmorphiated sufferers. But even after we have done all the narrowing down that subjects' reports permit, there will inevitably remain more than one candidate physical state P for any given phenomenal state C.

## 2.9 **Structure and substance**

The easiest way to see this is to take up a point mentioned by Davies at various points. Not all contemporary type-B physicalists aim to identify conscious states with strictly physical states. For there is also the option of identifying conscious states with properties that supervene on such strictly physical states, such as functionalist states defined in terms of causal structure. Thus one possibility might be to identify pain with the strictly physical state P which distinguishes those humans who are in pain from those who aren't. In this case pain would be identified with the property of having such-and-such organic molecules arranged into neurons in such-and-such ways. But then there is the option of abstracting away from the specific way this arrangement is realized in humans, and identifying pain with the structural property S of having such and such a causal arrangement of 'neurons' without specifying what the 'neurons' are made of.

Now these are certainly different accounts of the nature of pain. To see this clearly, consider the possibility of a 'silicon doppelgänger'—a being whose brain is structured just like yours, down to a fine level of detail, but whose 'neurons' are made of silicon-based circuits rather than organic molecules. If pain is identical with the strictly physical human property P, then this being will not experience pain. On the other hand, if pain is identical with the structural property S, then the doppelgänger will have pain.

The problem should now be becoming clear. How are we to decide between the claim that pain is identical to the strictly physical state P and the claim that

it is identical to the structural state S? The trouble is that both these states will be present in all humans who report pain, and both absent from all who deny it. After all, the two states are coextensive within humans, so there is no way that they can be distinguished by empirical research that relies on human reports of pain (cf. Levine 1983; Block 2002).

You might think that this is just a familiar problem of insufficiently varied data. Don't we simply need some cases where the two properties dissociate? A silicon doppelgänger, for example, would help. This would give us a case with the structural S but not the strictly physical P. So in principle shouldn't we be able to decide the issue by empirically testing whether a silicon doppelgänger is in pain or not?

However we can see that, even if we did have such a test case, it wouldn't really help. We already know that the doppelgängerwill say 'I am in pain' when it has S. After all, it is structured just like us. But it would say this even if conscious pain did depend on the organic realization P and were absent in beings that have S without P. (In that case, we could think of the doppelgängeras using the term 'pain' for a different state from our conscious pain.)

The difficulty of deciding between structural and strictly physical properties as the material essence of conscious properties is just one of a number of analogous problems that arise for consciousness research (see Papineau 2002, section 7.7). Thus, to take just one further example, consider the difficulty of deciding whether pain consists of (a) some material state M that is present just in case humans report pain or (b) M-plus-the-brain-processes-underlying-reports-of-pain (call this latter state 'H'—for 'higher-order'). Here too it seems clear that empirical research will be unable to adjudicate between M and H, since the two will be coextensive in humans: by definition, H will always be accompanied by M; and M must always be accompanied by H (brain processes underlying pain reports), given that M is picked out in the first instance as a state that is present whenever humans report pain (cf. Block, in press).

In this case too it is clear that the problem won't simply be resolved by gathering more varied data. We can of course seek out beings who have M but who lack the higher-order mental processes required for H. Vervet monkeys may well fit this bill. But such cases won't decide whether M without H suffices for pain. Since the subjects involved will lack the mental wherewithal to report on whether or not they feel pain, they will cast no light on the issue.

## 2.10 **Epistemology or semantics?**

What are we to make of these methodological conundrums? One possible response is that there is some kind of epistemological barrier preventing us

from finding out about the material nature of conscious states—there really is a fact of the matter about whether silicon doppelgängers or vervet monkeys feel pain, but we human investigators are for some reason prevented from ascertaining it.

I'm not happy with this conclusion. What exactly is supposed to be stopping us from finding out these putative facts? It's not as if the relevant data are far away, or too small for our instruments too detect, or only manifest at high energies that are very expensive to engineer. But in the absence of any such specific difficulty it seems mysterious that there should be some barrier to our finding out about the material nature of conscious states.

I prefer a different kind of explanation. The difficulty isn't that we can't access the facts, but that there aren't any real facts here to start with. The phenomenal concepts we use to pose the relevant questions—for example, do silicon doppelgänger s or vervet monkeys have pain?—aren't sharp enough to determine definite answers. The phenomenal concept pain is fine for distinguishing those humans who are in pain from those who aren't. But it goes fuzzy when we try to extend it to further cases. Asking whether a quite different kind of being feels pain is like asking whether it is five o'clock on the sun right now.

This suggestion might seem ad hoc. It is not in general a good idea to conclude that there is no fact of the matter just because we can't find out about something. However in this case there are good independent reasons for thinking that phenomenal concepts are insufficiently precise to determine the relevant facts. Phenomenal concepts are peculiar, imagistic, introspective ways of thinking about different kinds of conscious states. They serve a perfectly good purpose in enabling us to keep track of the mental lives of other human beings. But there is no reason to suppose that they draw a sharp line right through nature.

## 2.11 **The intuition of distinctness again**

At first sight it might seem absurd to deny that there is a fact of the matter about whether different kinds of being feel pain or other conscious states. Surely it is either like this for them, or it isn't. How can it be indeterminate whether the experience I am undergoing is also present in them?

But I wonder whether this intuitive reaction to the suggestion of indeterminacy is not just another manifestation of the intuition of distinctness I discussed earlier. Suppose that you think that certain brain processes give rise to the presence of some extra non-physical conscious mind-stuff. Then you will certainly think that there is some fact of the matter as to whether or not this same mind-stuff is present in different kinds of being. On the other hand, it is not

so clear that this conviction of determinacy will remain if you fully believe that conscious states are nothing but physical processes.

In support of this diagnosis, note that when we think of mental categories in non-phenomenal terms, there seems nothing intuitively objectionable about the idea that their application might sometimes be indeterminate. For example, the human visual system uses non-conscious stereopsis to help judge distance: it matches features across the left and right eye images to estimate the distance of their common source. Suppose some other beings also judge distance by matching features, but use different features from humans—relative light intensity rather than edges and lines, say. Or suppose they match the same features, but the processing is realized in silicon-based circuits rather than organic neurons. Is this still stereopsis? The natural reaction here is—who cares? Why should the concept of stereopsis be refined enough to determine a definite answer to these questions? If we do need a definite answer, we can refine the concept to determine one, but that's clearly not the same as ascertaining whether the divergent beings really instantiate our prior notion of stereopsis.

I infer from this comparison that our resistance to suggestions of indeterminacy arises specifically in those cases where we are thinking of mental states in phenomenal terms—in terms of what they are like, so to speak. When we do think of our mental life in this phenomenal way, then we are hard-pressed to avoid the intuitive thought that it involves some ontological addition to the physical processes of the brain. And once we slip into this dualist frame of mind, we will of course think there are definite facts about where the extra mind-stuff appears. But all this is driven by the dualist intuition that is incompatible with genuine physicalism.

Consider this analogy (cf. Papineau 2003). Eskimos make extensive use of whale oil, for heating, lubrication and many other purposes. At some point in Eskimo history, commercial manufacturers introduced a substitute petroleum product which behaved just like natural whale oil. Did the original Eskimo notion (let's call it '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, then it did.

Of course there was likely to be no answer here. Nothing in previous Eskimo thinking need have determined which way 'whale oil' should be understood when this question arose. Of course, the question could have been resolved by refining the Eskimo concept. The Eskimos could have decided henceforth to understand 'whale oil' as referring to a biological or chemical substance, in which case the new stuff wouldn't have counted as 'whale oil'. Alternatively, they

could have decided to understand it as referring to anything with the appropriate appearance, in which case the new stuff would have counted as 'whale oil'. But clearly neither option need have been determined by the prior Eskimo notion of 'whale oil'.

But now imagine people who were unhappy about the idea that this was just a matter for decision. They said, 'Yes, I can see that this new manufactured stuff is like the old stuff in appearance, while different in chemical constitution and biological origin. But that doesn't yet answer the important question, which is whether it is like the old stuff in being made of whale oil.'

Now clearly these people would have been assuming that being made of whale oil is some extra property, over and above constitution and behaviour. They would have been dualists about the property of being made of whale oil. Because of this, they would have thought that there must be some fact of the matter as to whether the new product involved this extra dualist whale oil or not.

I think the same about people who are convinced there must be some fact of the matter as to whether the doppelgänger is really in pain. They are assuming that pain is some extra property, over and above physical and structural properties, and so of course they think that there is a substantial matter as to whether this extra property is present in the doppelganger. But in truth there is no such extra property. There is nothing to refer to except the relevant physical and structural properties, and our phenomenal concepts leave it indeterminate which of these is to be identified with pain.

## 2.12 **Conclusion**

I am not aiming to belittle empirical research into consciousness. As I said, it can tell us many interesting and surprising things about the material processes present in human beings in different conscious states. However, there are certain apparent questions about consciousness that empirical research won't be able answer. For it will always leave us with a plurality of candidates for the material essence of any conscious property. And in the face of this plurality empirical science will be impotent, for the methodology of consciousness research offers no handle by which to separate material properties that are perfectly correlated in normal humans.

To this extent, scientific research into consciousness is fated to deliver less than we might have hoped. 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 precisely 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. (2002). The harder problem of consciousness. *Journal of Philosophy* 99, 391–425.

Block, N. In press. consciousness, accessibility, and the mesh between psychology and neuroscience. *Brain and Behavioral Sciences*.

Bloom, P. (2004). *Descartes' Baby*. New York: Basic Books.

Chalmers, D. and Jackson, F. (2001). Conceptual analysis and reductive explanation. *Philosophical Review* 110, 315–360.

Dennett, D. (1978). Why you can't make a computer that feels pain. In *Brainstorms*, pp. 190–229. Cambridge, MA: Bradford Books.

Goodale, M. and Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences* 15, 20–25.

Kripke, S. (1980). Naming and necessity. Oxford: Blackwell.

Levine, J. (1983). Materialism and qualia: the explanatory gap. *Pacific Philosophical Quarterly* 64, 357–358.

Libet, B. (1993). The neural time factor in conscious and unconscious events. In Bock, G. and Marsh, J. (eds) *Experimental and Theoretical Studies of Consciousness*, pp. 122–137. London: Wiley.

Melnyk, A. (2003). Papineau on the Intuition of Distinctness. SWIF Forum on *Thinking about Consciousness* http://lgxserver.uniba.it/lei/mind/forums/004_0003.htm.

Papineau, D. (1993). Physicalism, consciousness, and the antipathetic fallacy. *Australasian Journal of Philosophy* 71, 169–183.

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

Papineau, D. (2003). Could there be a science of consciousness? *Philosophical Issues* 13, 205–220.

Papineau, D. (2006). Comments on Strawson's 'Realistic monism: why physicalism entails panpsychism'. *Journal of Consciousness Studies* 13, 100–109.

Papineau, D. (2007). Kripke's proof is *ad hominem* not two-dimensional. *Philosophical Perspectives* 21, 475–494.

Weiskrantz, L. (1986). *Blindsight.* Oxford: Oxford University Press.