The Content of Physicalism

by

Daniel Stoljar
BA (Hons.), University of Sydney

Submitted to the Department of Linguistics and Philosophy in
Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the
Massachusetts Institute of Technology
September 1995

© 1995 Daniel Stoljar. All rights reserved.

The author hereby grants MIT permission to reproduce and to distribute
publicly paper and electronic copies of this thesis document in
whole or in part.

Signature of Author: ........................................... ........................

Department of Linguistics and Philosophy

Certified by: .............................................................

Ned Block
Professor, Department of Linguistics and Philosophy
Thesis Supervisor

Accepted by: ..............................................................

George Boolos
Chairman, Departmental Graduate Committee
Department of Linguistics and Philosophy
The Content of Physicalism
by
Daniel Stoljar

Submitted to the Department of Linguistics and Philosophy in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Philosophy

Abstract

Many philosophers of mind are concerned to defend the thesis called physicalism (or materialism); many others are concerned to refute it. Nevertheless, there is no generally agreed on idea of what physicalism is, and why it should matter whether the mental is physical. My thesis consists of four essays whose concern is with what physicalism is in its most plausible version, and what the importance of the thesis might be for the philosophy of mind.

I begin with the question of whether it is possible to hold physicalism in a nonreductive form. Many philosophers are attracted to nonreductive physicalism because they think that considerations of multiple realizability, and to a lesser extent Davidsonian considerations about the anomalism of the mental, establish that it is impossible to hold physicalism in a reductive form. Nonreductive physicalism is a doctrine that divides into two parts. The first part -- the physicalist part -- is that mental properties and facts supervene on physical properties and facts. The second part -- the nonreductivist part -- is that this supervenience thesis does not imply that mental properties and facts are reduced to physical properties and facts. In chapter one, I argue that, when properly understood, these two claims are not cotenable. Supervenience implies that there are necessary connections between mental and physical properties. However, there is no way to account for these necessary connections without supposing that the supervenient mental properties are reduced to physical properties. Hence supervenience implies reduction.

In chapter two, I consider a different way of developing nonreductive physicalism. Stephen Schiffer argues that physicalism can be combined with the view that there are irreducible mental properties if one adopts a nominalist or deflationary view of properties in general. I reply that considerations from the metaphysics of properties cannot save nonreductive physicalism: supervenience implies reduction even if one adopts the view of properties Schiffer recommends.

My discussion of nonreductive physicalism has a conditional conclusion: nonreductive physicalism reduces to reductive physicalism, and thus, if one wants to be a physicalist, one must be a reductive physicalist. Accordingly, in chapter three, I turn to reductive physicalism: the thesis that mental properties
simply are physical properties. Jaegwon Kim argues that reductive physicalism entails that there can be no psychological laws or kinds. I reply that while some versions of reductive physicalism certainly have this result, the best version of the doctrine does not. According to this best version, mental properties are second order properties. However, I also argue that even this best version of reductive physicalism faces a serious objection: it could only be true on the condition that we revise our current conception of what a mental property is. Revising our conception of what a mental property is seems a very high price to pay for physicalism, so before doing that, we had better assess the reasons for believing physicalism.

In the final chapter, then, I assess the reasons for believing physicalism. Physicalists have a number of different reasons for believing their doctrine, but perhaps the most widespread is the idea that one should believe physicalism because to renounce it would be, as Schiffer has put it, "to renounce the scruples of the natural scientist". I reply that this is a mistake born from missing an ambiguity in the word 'physical'. In one sense, 'physical' applies to a property or fact which the empirical sciences are in the best position to describe. In another sense, 'physical' applies to a property or fact which a certain sub-class of the empirical sciences are in the best position to describe, namely, the subclass which includes contemporary physics, chemistry, biology and neuroscience. What is at issue in the philosophy of mind is physicalism where 'physical' has this second stricter sense. However, under that interpretation, to renounce physicalism is not, I argue, to renounce the scruples of the natural scientist. Moreover, under that interpretation, there is no strong pressure to revise our conception of what a mental property is in order to ensure the truth of physicalism.

Supervisor: Ned Block
Title: Professor of Philosophy
Acknowledgments

I owe an embarrassingly large debt to my thesis committee -- Ned Block, Judith Jarvis Thomson and Robert Stalnaker -- for discussion, criticism, and encouragement. I am also deeply grateful to Noam Chomsky for comments and discussion and for first asking me what on earth I thought physicalism was anyway. In addition, I would like to thank the following for their help: Andrew Botterell, Sylvain Bromberger, Alex Byrne, Leonard Clapp, Josh Cohen, Michael Glanzberg, Tracy Isaacs, Ned Hall, Tim Hinton, Paul Horwich, David Hunter, Robert Kermode, Joe Lau, Cathy Legg, Jennifer Noonan, Simon Saunders, Robert Stainton, Jason Stanley and Zoltan Szabo.

The title of my dissertation is borrowed from a three page paper by J.J.C.Smart, published in 1978.
Table of Contents

Abstract 2
Acknowledgments 4

Chapter One: Reductive and Nonreductive Physicalism 7
Chapter Two: Nominalist Physicalism 39
Chapter Three: Disjunctive and Second Order Physicalism 69
Chapter Four: 'Physical' and Physicalism. 97

Bibliography 124
There is a war
Between the ones who say
There is a war
And the ones who say
That there isn’t.

Leonard Cohen
Many philosophers nowadays agree that physicalism in the philosophy of mind is to be formulated as a supervenience thesis: according to physicalism, mental properties supervene on physical properties. After that, however, the level of agreement drops off. One source of dispute concerns the fact that there are various non-equivalent supervenience theses, and hence that to say that mental properties supervene on physical properties is not yet to say anything very precise. Another source of dispute concerns the issue of reductionism. Some philosophers agree with David Lewis that "A supervenience thesis is, in a broad sense, reductionist" (1983: 358), and therefore that any physicalism formulated in terms of supervenience is reductive physicalism. For many others, however, the whole point of formulating physicalism in terms of supervenience is that supervenience is, in a broad sense, nonreductionist, and therefore that any physicalism formulated in terms of supervenience is -- or might be -- nonreductive physicalism.

These disputes are of course not unrelated. Just as there are various non-equivalent supervenience theses, there are various non-equivalent reductionist theses. It is therefore very plausible to suppose that the claim that supervenience theses are reductionist will be true under some interpretations and false under others. More generally, one begins to suspect that the debate in philosophy of mind about whether it is possible to be a nonreductive physicalist is driven solely by equivocation on its central terms. Both sides, it is easy to suppose, might be perfectly correct, so long as it is granted that each
can operate with its own definitions of 'physicalism', 'supervenience', and 'reductionism'.

My aim in this paper is to demonstrate that while there is a certain amount of truth to this suspicion, it is not the whole truth. More particularly, I will argue, first, that on a proper understanding of the notions involved, there is a genuine question of whether nonreductive physicalism is possible, and second, that this genuine question is properly answered in the negative.

The paper falls into two halves. In the first half, I consider the kind of supervenience thesis that physicalists should find attractive, concentrating mainly on some recent remarks of Jaegwon Kim. In the papers collected in *Mind and Supervenience*, Kim distinguishes two sorts of supervenience claim, weak supervenience and strong supervenience, and then goes on to argue that physicalism is best formulated as a strong, rather than a weak, supervenience thesis. I agree that physicalism is not best formulated as weak supervenience, but I will argue that physicalism is not best formulated as strong supervenience either. In place of strong supervenience, I propose another kind of supervenience thesis, which I call weak $^2$ supervenience. It is to weak $^2$ supervenience, rather than strong supervenience, that physicalists should look in order to formulate their doctrine.

In the second half of the paper, I go on to consider whether physicalism so defined is consistent with the denial of reductionism. While there are a number of different versions of the thesis of nonreductive physicalism, I argue

---

1 In addition to weak and strong supervenience, Kim discusses and rejects a third sort of supervenience claim, global supervenience, as well as the corresponding idea that physicalism is the doctrine that the mental globally supervenes on the physical. See Kim 1993: Chapters 4 and 5. Global supervenience theses are interesting; and under some interpretations would avoid the problem I will raise against strong supervenience thesis. In this paper, however, I will limit myself to what are sometimes called 'co-instantiation' theses, i.e. strong supervenience theses and weakenings of strong supervenience theses.
that there is only one interesting version, and this is the version associated with Donald Davidson. I further argue, however, that this interesting version of nonreductive physicalism is mistaken, and thus that to this extent physicalism does indeed entail reductionism.

II
Let us suppose that M is the family of mental properties and P is the family of physical properties. The claim that mental properties weakly supervene on physical properties can then be stated as follows:

(WS) Necessarily for any object x, and any property F in M, if x has F, then there exists a property G in P such that x has G, and if any y has G it has F.

Kim argues persuasively against the identification of physicalism and (WS) (1993: 63). Central to physicalism is the idea that the physical facts about a person entail the mental facts about the person. However, (WS) does not secure this entailment. For it is quite consistent with (WS) that two people in different possible worlds are physically identical and yet differ mentally. But if two people in different possible worlds can be identical physically and yet differ mentally, then the physical facts about a person do not entail the mental facts. Hence (WS) is not a formulation of physicalism.

If (WS) is not physicalism, another possibility is that physicalism is the thesis that mental properties strongly supervene on physical properties. This latter claim can be stated as follows:

(SS) Necessarily, for any object x and for any property F in M, if x has F, then there exists some property G in P such that x has G, and, necessarily, if any y has G, it has F.
Kim argues that physicalism ought to be stated as (SS), and certainly this claim is more plausible than the parallel claim about (WS).2 For one thing, (SS) secures, where (WS) does not, the entailment of mental facts by physical facts. If (SS) is true, in other words, it is impossible for two people in different possible worlds to be physically identical and yet mentally different.

Even if (SS) is more plausible as a statement of physicalism than (WS) however, there is nonetheless a serious problem with the suggestion that (SS) is the statement of physicalism. The problem is that (SS) excludes possibilities that many physicalists think are genuine possibilities. For example, many physicalists want to insist that the relation between mental and physical is a contingent relation in the following sense: While it is as a matter of fact true that mental properties supervene on physical properties, it is also possible that a mental property might be instantiated without its being co-instantiated with a physical property. In other words, such physicalists want to accept a thesis we might call the contingency thesis: the thesis that one might have a mental property without having any physical properties. However, (SS) is inconsistent with the contingency thesis. Hence physicalists who want to accept the contingency thesis must reject (SS) as a statement of their position, and (SS) cannot be the statement of physicalism.

Why is (SS) inconsistent with the contingency thesis, i.e., why does it exclude the possibility of having a mental property without also having a physical property? The easiest way to see this is by reflecting on an idea that figures prominently in Kim's discussion of strong supervenience: the idea that (SS) entails the necessary coextension of mental and physical properties. According to strong supervenience, Kim says:

---

2 Kim's suggestion that (SS) is the statement of physicalism is very popular one. For recent endorsements, see Horgan (1994) and Yablo (1992).
Whenever a supervening property P is instantiated by an object, there is a subvenient property Q such that the instantiating object has it and the following conditional holds: necessarily, if anything as Q, then it has P. So the picture we have is that for the supervenient property P, there is a set of properties, Q₁, Q₂,...in the subvenient set such that each Qᵢ is necessarily sufficient for P. Assume this list contains all the subvenient properties each of which is sufficient for P. Consider their disjunction: Q₁ or Q₂ or...(or UQᵢ for short). It is easy to see that this disjunction is necessarily coextensive with P. First, it is clear enough that UQᵢ entails P, since each disjunct does. Second, does P entail UQᵢ? Suppose not: something then, say b, has P but not UQᵢ. According to [strong supervenience], b has some property in the subvenient set, say S, such that necessarily whatever has S also has P. But then S must be one of the Qᵢ, and since b has S, b must have UQᵢ. So P entails UQᵢ. So P and UQᵢ are necessarily co-extensive [...]. (1993: 151-2).³

Now Kim is interested in necessary co-extension primarily because of reductionism. If (SS) entails the necessary co-extension of mental properties and physical properties, he argues, it also entails that mental properties are reduced to physical properties. I will return to the issue of reductionism later on, but for the moment I simply want to point out the following: Whatever is the case with reductionism, it is certainly true that if (SS) entails the necessary co-extension of mental and physical properties, then (SS) is inconsistent with the contingency thesis.

³ For further details of the argument, see Kim 1993: 151-2 and for discussion, see Horgan (1994: 576). For the argument in a somewhat more complicated form, see Kim 1993: 70-1, and for discussion, see Van Cleve 1990 and Bacon 1990.
For consider a particular mental property -- the property of being in pain, as it might be. If (SS) is true, and if Kim's argument in the quoted passage is sound, then there is some physical property, call it P*, which is necessarily co-extensive with the property of being in pain. Moreover, if physicalism is (SS), then physicalism entails that P* is necessarily co-extensive with being in pain. But this means that physicalism cannot allow for contingency in the sense we have been discussing: if being in pain is necessarily co-extensive with a physical property, it is impossible to be in pain and lack that physical property. \textit{A fortiori}, it is impossible to be in pain and lack any physical property.

III

Our argument so far has been that (SS) cannot be the statement of physicalism because there are physicalists who must rejected it, viz., those physicalists who accept the contingency thesis, and therefore accept the possibility that one might be in pain, for example, and have no physical properties. However, since the idea that (SS) is physicalism is a very popular view, before we reject it we should consider how a friend of (SS) might respond.

\textit{Response #1: Physicalism inconsistent with Contingency.}

To begin with, one might argue that, while there is a logical possibility of having a mental property without also having a physical property, this is not a possibility that physicalists can acknowledge. In other words, one might argue that physicalists of any stripe must reject the contingency thesis, and therefore that it is no criticism of the idea that (SS) is the statement of physicalism that (SS) is inconsistent with the thesis.

However, it is simply not true that physicalists of any stripe must reject the possibility of someone's having a mental property without also having a
physical property. According to one plausible and influential philosophical tradition, for example, mental properties are to be identified with topic-neutral properties. According to this kind of view, the property of being in pain is identical to the (second order) property of having some property with a certain causal role. Now on the one hand, this position is clearly consistent with physicalism, because it is consistent with the topic-neutral account of mental properties that all the properties which actually play the causal role definitive of being in pain are physical properties. On the other hand, however, the topic-neutral account is also consistent with the contingency thesis. After all, it is perfectly possible that the property that has the causal role definitive of being in pain is a non-physical property.

Response #2: Contingency Thesis False.

Second, one might try to respond to our argument by denying the contingency thesis and thereby denying that it is possible that one could be in pain without having a physical property. Of course, if the contingency thesis is false, then the fact that (SS) provides no place for it is no objection at all to the idea that (SS) is physicalism.

---

4The tradition I have in mind here includes the materialism of Smart (1959), Armstrong (1968) and Lewis (1966), as well as most contemporary forms of functionalism. (For discussion of functionalism see Block 1980, 1990.) The idea of topic neutral language (and of topic neutral properties) is due to Smart 1959. Smart wrote: 'When a person says "I see a yellowish-orange after image", he is saying something like this: "there is something going on in me that is like going on when I have my eyes open, am awake, and there is an orange illuminated in good light in front of me, that is, when I really see an orange' (1959: 146). Lewis later captured the topic-neutrality of Smart's account by saying that "It is not a materialist principle, nor does it ascribe materialism to whoever speaks of experiences. Rather it is an account of the parlance common to all who believe that experiences are efficacious outside their own realm. It is neutral between theories -- or lack of any theory about what sort of real and efficacious things experiences are: neural states and the like, pulsations of the ectoplasm and the like, or just experiences and nothing else "(1966: 102).
However, the contingency thesis is more difficult to deny than one might imagine.\(^5\) Perhaps from a certain point of view it can seem outlandish to suppose that someone might have a mental property without having any physical property. However, any outlandishness in this idea can be almost wholly dispelled if we examine more closely what marks a property as a physical property. For most philosophers, a physical property is a property that figures in the contemporary physical sciences, that is, contemporary physics, chemistry, biology and neuroscience.\(^6\) However, if this is the notion of a physical property that is at issue, then it seems clear enough that one might have a mental property without having a physical property. After all, it is not hard to imagine oneself being in pain, and yet it also being false that one has any property discussed in contemporary physics, chemistry, biology and neuroscience.

Might it not be argued on behalf of (SS), however, that while it certainly appears possible that one might have a mental property without any physical property, this appearance is in fact mistaken? In particular, might a friend of (SS) not explain away the appearance that the contingency thesis is true? The problem with this suggestion, however, is that any attempt to explain away the appearance that it is possible to have a mental property without having a

\(^5\) It is worthy of note that, in this regard, physicalism is quite unlike parallel claims in meta-ethics. Many philosophers are prepared to defend the thesis that moral facts and properties supervene on natural facts and properties. However, there is little or no temptation to think that such theses are contingent. Since, as is well known, the notion of supervenience entered the philosophy of mind via meta-ethics, it is tempting to speculate that many philosophers were so taken with the analogy between the relationship between the moral and the natural on the one hand, and the relationship between the mental and the physical on the other, that they simply overlooked disanalogies between the cases, of which the contingency thesis is a prime example.

\(^6\) For a recent statement, see Jackson 1994.
physical property would quickly land the friend of (SS) in a series of problems made clear by Saul Kripke.

As Kripke has famously argued, if two properties A and B are necessarily co-instantiated, then the only way to explain away the appearance that one can have A without B is by producing a third property, C, which on the one hand is superficially very much like A, but on the other hand is not necessarily co-instantiated with B. Now, according to (SS), the property of being in pain is necessarily co-instantiated with a physical property P*. It follows then, that in order to explain away the appearance that one could be in pain and not have P*, it would be necessary to produce another property which on the one hand is superficially very much like being in pain but which one the other hand is only contingently associated with P*, i.e., one could have this other property without having P*. However, as Kripke has even more famously argued, a property that is superficially very much like being in pain is plausibly regarded as being in pain (1980: 151). And this means that the friend of (SS) will not succeed in explaining away the possibility that one might be in pain and yet not have any physical property.

Response #3: Closure Assumptions Rejected.

If the contingency thesis cannot be rejected, another possibility is that one might avoid our argument by interpreting (SS) so that it is consistent with the

---

7 Kripke was primarily interested in claims of identity rather than claims of necessary co-instantiation, but the point remains the same. He wrote: "Any necessary truth, whether apriori or a posteriori, could not have turned out otherwise. In the case of some necessary a posteriori truths, however, we can say that under appropriately qualitatively identical evidential situations an appropriate corresponding qualitative statement might have been false" (1980: 142).

8 It is important to note here that, unlike Kripke, I am not arguing that physicalism is false. I am merely arguing that physicalism would be false, were it identical with (SS).
thesis. One way one might try to do this, for example, is by denying the assumption about properties that is assumed in Kim's argument that (SS) entails the necessary co-extension of mental and physical properties. In the quotation above, Kim is clearly assuming that the family of physical properties is closed under disjunction: that is, he assumes that a disjunction of physical properties is itself a property, and then argues that this disjunctive property is necessarily co-extensive with (to continue with our example) the property of being in pain. But of course one might deny this closure assumption, and thereby deny that (SS) entails there is a physical property which is co-extensive with the property of being in pain.

However, while it is quite true that the assumption of disjunctive closure is controversial, I adopted it only for simplicity, and not because it is required for our purposes. Certainly Kim's argument about necessary co-extension requires that the physical family of properties is closed under disjunction. However, all that is required for our argument that (SS) is inconsistent with the contingency thesis is the weaker claim that, necessarily, for every mental property F and for any object x, if x lies in the extension of F, then there is some physical property G such that x has G. This weaker claim follows from (SS) even without the closure assumption. But even this weaker claim rules out the contingency thesis since it rules out the possibility that something could be in pain without having a physical property.

Response #4: Modal Operators Restricted.

Finally, one might try to restrict the modal operators that appear in the definition of (SS). So far, I have been assuming that the modal operators in both (WS) and (SS) are to be interpreted without restriction, viz., as applying relative to all possible worlds. However, it might be possible to interpret them
as applying relative only to a certain class of *nomologically* possible worlds, rather than to all possible worlds. This maneuver would render (SS) consistent with the contingency thesis, since even if physicalism is true in all worlds of a certain class, it might nevertheless be false in worlds which fall outside that class.

This suggestion certainly has some plausibility; in particular, if one is intent on articulating physicalism as a strong supervenience thesis, it is hard to see how one might account for the contingency thesis without restricting the modal operators in something like the way suggested. Unfortunately however, there is a quite independent problem with this restriction proposal, and this rules it out as a way of answering our problem about contingency. 9

Let us suppose that (SS) is interpreted restrictively so that it is true only at a certain class of nomologically possible worlds. The problem with the identification of physicalism and (SS) so understood is that physicalism is now a much stronger thesis than it is typically taken to be. In particular, physicalism so understood would be false in a world that had no *irreducible* mental facts or properties not as a matter of law, but merely accidentally. But there seems no reason to think that physicalism would be false in such a world. After all, there is some plausibility to thought that, in our world, while there are no mental forces on a par with physical forces, the conservation of energy laws do not *rule out* mental forces. But that is not a reason for thinking that physicalism is false in our world.

In summary then, there seems to be no way that one can defend (SS) from the charge that it is inconsistent with the contingency thesis. On the one hand, this looks to be a thesis that physicalists cannot reasonably give up. On

---

9 This problem is discussed in Horgan (1982) and in Lewis (1983).
the other hand, however, there is no way to interpret (SS) so as it is consistent with the thesis.

IV

To this point our discussion of the relation between physicalism and supervenience has been wholly critical. We have examined Kim's suggestion that physicalism ought to be stated as a strong supervenience thesis and found it wanting on the grounds that strong supervenience is too strong: strong supervenience is inconsistent with the contingency thesis, and this thesis is something that physicalists cannot give up.

If physicalism is not strong supervenience, however, then what kind of supervenience thesis is it? The issues here are of course extremely complicated; however, I think that what has been said so far enables us to make a number of suggestions about the kind of supervenience thesis that is likely to recommend itself to physicalists.

To begin with, there is no point reconsidering the idea that physicalism correctly stated is a weak supervenience thesis. This is not only because Kim's rejection of (WS) as a statement of physicalism is persuasive; it is also because the problem we have raised for strong supervenience is equally a problem for weak supervenience. For weak supervenience, no less than strong supervenience, is inconsistent with the contingency thesis, since it too entails that necessarily if something has a mental property, then there is some physical property it has. In other words, weak supervenience should be rejected as a statement of physicalism for our reasons as well as Kim's: weak supervenience in one sense is too weak, but in another sense it is too strong.

Even if physicalism properly construed is neither (WS) or (SS), however, there is a simple permutation of these theses which is, I think, worthy of
consideration. The difference between the two supervenience theses that Kim
discusses is that strong supervenience has a second (embedded) modal
operator that weak supervenience lacks. However, it is not this second operator
that is the cause of our problem over contingency. Since our argument applies
equally to weak and strong supervenience, the operator that distinguishes them
cannot be the source of the trouble.

But this suggests that one way to avoid our problem is by keeping the
second operator that is distinctive of strong supervenience, and dropping
rather the first operator. That would result in the following sort of
supervenience thesis, which we might call weak2 supervenience:

(W2S) For any x, and for any property F in M, if x has F, then there is
some property G in P such that x has G and, necessarily, for any
y, if y has G, then it has F.

For purposes of stating physicalism, (W2S) apparently avoids the problems of
both (WS) and (SS). Unlike (WS), it has the appropriate modal strength: (W2S)
generates an entailment of the mental facts about a person by the physical facts.
Unlike (SS), on the other hand, (W2S) allows for the contingency thesis; at least,
given the correct conditions of closure, (W2S) does not entail that every mental
property is necessarily co-extensive with some physical property. 10

In suggesting that (W2S) might be the right kind of supervenience thesis
for physicalists I do not mean to suggest that it is the only kind that might do
the job. Indeed, it seems to me that our previous discussion suggests quite
generally that what physicalists require of a supervenience thesis is that it

---

10 The qualification about closure conditions is important here. In particular, if (W2S) is
interpreted so as the physical family of properties is closed under negation, then (W2S) will
entail necessary co-extension just as (SS) does. I think the proper response to this problem is to
deny that the physical family is closed under negation, but I cannot defend that response here.
(I am indebted here to Robert Stalnaker.)
satisfies the following two conditions: the thesis must (a) guarantee the entailment of mental facts by physical facts and (b) not guarantee the reverse entailment. Weak$_2$ supervenience satisfies both conditions, and to that extent it ought to be interesting to physicalists. But there are certainly other kinds of supervenience theses that also satisfy the conditions; it may even be that (W$_2$S) is not the most attractive supervenience thesis that satisfies the conditions. What makes (W$_2$S) interesting in the present context is mainly that it is a fairly obvious (but to my knowledge unnoticed) variation of the supervenience theses that Kim discusses.

V

So far then, our focus has been on the kind of supervenience theses that physicalists should find attractive. We have considered various kinds of supervenience theses and suggested that physicalism is best formulated as a kind of supervenience thesis of which weak$_2$ supervenience is an instance. Weak supervenience theses are too weak since they do not guarantee the entailment of mental facts by physical facts; but strong supervenience theses are too strong since they do not allow for the idea that physicalism is consistent with the contingency thesis.

I turn now to the issue of reductionism. As I noted at the beginning, for many philosophers the whole point of formulating physicalism in terms of supervenience is that supervenience is, in a broad sense, nonreductionist, and therefore that any physicalism formulated in terms of supervenience is -- or might be -- nonreductive physicalism. Indeed I think it is fair to say that rather more philosophers of mind conceive of themselves as nonreductive physicalists than as reductive physicalists, and that to this extent Jaegwon Kim
is correct when he says that nonreductive physicalism is the "received view" (1993: 339) in the philosophy of mind.

On the other hand, however, it is none too clear what precisely the received view is.

Of course, it is easy enough to discern the broad contours of the position. In broad outline, nonreductive physicalism is a doctrine that divides into two parts. The first part -- the physicalist part -- is that mental properties supervene on physical properties. The second part -- the nonreductionist part -- is that mental properties are not reduced to physical properties. Moreover, as we have just seen, it is possible to bring the first part of the doctrine into sharper relief by clarifying what the supervenience relation is that, according to nonreductive physicalism, holds between mental and physical properties. In particular, if what I have argued so far is correct, the supervenience relation required by physicalists is weak$_2$ supervenience. Nevertheless, the second part of nonreductive physicalism remains murky. After all, what does it mean to say that mental properties are reduced to physical properties, and, more to the point, what does it mean to deny that they are?

Now of course there are many things that it might mean to deny that mental properties are reduced to physical properties. Fortunately, however, our purposes here do no require discussion of them all. For in fact it seems to me that, while there are many different theses of reductionism, there are three theses which have achieved prominence in the literature on nonreductive physicalism. The thing for us to do, therefore, is to distinguish among these three reductionist theses, and on the basis of this distinction, to define three versions of nonreductive physicalism.

According to the first view -- which we might call reductionism$_1$ -- reductionism is the thesis that mental properties are identical either with
physical properties or with topic-neutral properties, where by 'physical
property' I mean a property that figures in contemporary physics, chemistry,
biology and neuroscience, and by 'topic neutral property' I mean the kind of
second order property I discussed earlier, viz., the property of having some
property that has a certain causal role. According to the second view -- which
we might call reductionism$_2$ -- reductionism is the thesis that mental properties
are identical with physical properties only. And according to the third view --
reductionism$_3$ -- reductionism as a doctrine that says nothing at all about
property identity. Rather, it is the thesis that there are psychophysical laws
linking the instantiation of mental properties and the instantiation of physical
properties.

I have no quarrel with any of these conceptions of reductionism; at any
rate I certainly do not want to insist that one of the three -- reductionism$_1$, as it
might be -- is the proper notion, or even that there is such a thing as the proper
notion of reductionism. Nevertheless, it is important to point out that we are
here in the presence of three rather different notions of reductionism.
Reductionism$_2$, for example, is clearly different from reductionism$_1$, since the
idea that mental properties are identical with topic neutral properties is a
reductionist idea according to reductionism$_1$, but it is a non-reductionist idea
according to reductionism$_2$. Similarly, reductionism$_3$ is obviously different
from reductionism$_1$, since it seems clear enough that one might identify mental
properties with topic neutral properties without thinking that there are
psychophysical laws linking mental and physical properties. And there are
other differences besides.$^{11}$

$^{11}$ It will not be important for us to discuss all the similarities and differences among the
different notions of reductionism.
Moreover, it is important to point out that these three different theses of reductionism generate three different versions of nonreductive physicalism. The first version -- we might call it \textit{nonreductive}_1 \textit{physicalism} -- is the version which combines physicalism with the denial of reductionism\textsubscript{1}. The second version -- \textit{nonreductive}_2 \textit{physicalism} -- is the version which combines physicalism with the denial of reductionism\textsubscript{2}. And the third version -- \textit{nonreductive}_3 \textit{physicalism} -- is the version which combines physicalism with the denial of reductionism\textsubscript{3}.

In short, then, while Kim is in one sense quite correct to say that nonreductive physicalism is the received view in philosophy of mind, his remark is also somewhat misleading. It is plausibly true that a statistical majority of philosophers of mind would describe themselves as nonreductive physicalists. On the other hand, however, it does not follow from this that the majority subscribe to the \textit{same} view. For, as we have just seen, there are (at least) \textit{three} versions of nonreductive physicalism! Hence Kim's remark is misleading: to say that nonreductive physicalism is \textit{the} received view imposes a false unification on what is in fact a rather disunited situation.

VI

Our first point concerning nonreductive physicalism, then, is simply that there are different versions of the doctrine, and therefore that it is misleading to suppose that nonreductive physicalism is a unified view. Perhaps most philosophers would describe themselves as writing under the banner of nonreductive physicalism -- but it is not clear that these philosophers have the \textit{same} banner in mind.

I want now to draw some consequences from our tripartite distinction among the varieties of nonreductive physicalism. The first consequence is
simply that the question of whether nonreductive physicalism is possible is rather more complicated that it at first appears. In particular, to raise the question of whether nonreductive physicalism is possible is really to raise three different questions, each of which might receive a different answer: the first concerns nonreductive\textsubscript{1} physicalism, the second nonreductive\textsubscript{2} physicalism, and the third nonreductive\textsubscript{3} physicalism.

The second consequence is that only one of these three questions is of any passing philosophical interest, and this is question about the possibility of nonreductive\textsubscript{1} physicalism.

For the fact is that if you are operating with either reductionism\textsubscript{2} or reductionism\textsubscript{3}, there is no question but that nonreductive physicalism is possible, i.e., there is no question but that you can be a physicalist and deny reductionism in these senses of the `reductionism'. Suppose, for example, that reductionism is the thesis that mental properties are identical to physical properties; that is, suppose that by `reductionism' you mean reductionism\textsubscript{2}. And suppose moreover that by `nonreductive physicalism' you mean nonreductive\textsubscript{2} physicalism, i.e., the version of nonreductive physicalism which combines physicalism and the denial of reductionism\textsubscript{2}. Then it is obvious that one can hold nonreductive physicalism. After all, as we saw in the previous section, one might be a physicalist and hold the topic neutral account of mental properties; but that account is inconsistent with reductionism\textsubscript{2}.

Similar remarks apply in the case of reductionism\textsubscript{3}, the thesis that there are psychophysical properties linking the instantiation of mental properties to the instantiation of physical properties. Suppose, for example, you were the kind of physicalist who held that mental properties are topic neutral second order properties. Then you could quite reasonably deny reductionism\textsubscript{3}, and therefore hold nonreductive\textsubscript{3} physicalism, since you might well deny that there
are no psychophysical laws linking the instantiation of second order properties with the first order physical properties that realize them. In other words, then, what goes for reductionism\textsubscript{2} goes also for reductionism\textsubscript{3}: any physicalism of the kind we discussed above -- that is, any physicalist who holds the topic neutral account of mental properties -- can quite consistently deny either reductionist thesis. It follows, that if by `nonreductive physicalism' you mean either nonreductive\textsubscript{2} physicalism or nonreductive\textsubscript{3} physicalism, then nonreductive physicalism is quite obviously possible.

On the other hand, however, the fact that nonreductive\textsubscript{2} physicalism and nonreductive\textsubscript{3} physicalism are consistent positions does not seem like a very exciting result. As we have seen, to qualify either as a nonreductive\textsubscript{2} physicalist or as a nonreductive\textsubscript{3} physicalist it is sufficient that one is a physicalist who holds the topic neutral account of mental properties. In the first place, however, it has been known for a very long time that one can be a physicalist and hold the topic neutral account of mental properties. If the much vaunted problem of whether one can be a nonreductive physicalist amounts only to the question of whether one hold physicalism together with the topic neutral account of mental properties, there doesn't seem like much of problem. In the second place, not only it is true that one can be physicalist and hold the topic neutral account of properties, it is also true that many philosophers who defend the topic neutral account for mental properties take themselves to be reductionist\textsuperscript{12} It follows, then, that there is no real difference between

\textsuperscript{12} Consider, for example, David Lewis's philosophy of mind, recently set out in `Reduction of Mind' (1994). Lewis clearly regards himself as a reductionist, as his title makes clear. On the other hand, Lewis also is committed to the idea that when people are in pain they have the topic neutral property of having some property with such and such a causal role. Indeed, Lewis even calls this property `being in pain' -- though he reserves the word `pain' for something else. This suggests that, as far as Lewis is concerned, it is reductionism\textsubscript{1} that matters. In short, Lewis is a reductionist in the sense of reductionism\textsubscript{1} but he is a non reductionist in the sense of reductionism\textsubscript{2}. 
nonreductive\textsubscript{2} physicalism and nonreductive\textsubscript{3} physicalism on the one hand, and this traditional version of \textit{reductive} physicalism on the other. In particular, the only difference between these versions of nonreductive physicalism and this traditional version of reductive physicalism is the disputed word 'reduction', and that, of course, is no difference at all.

Indeed, it is at this point in the discussion that one begins to suspect that there really is no issue about nonreductive physicalism, and that the debate about whether it is possible is driven solely by equivocation. For not only is it perfectly consistent to hold the conjunction of physicalism and the denial of either reductionism\textsubscript{2} or reductionism\textsubscript{3}, it is also true that many philosophers who describe themselves as nonreductive physicalists have only these positions in mind. In other words, many so-called nonreductive physicalists are in fact either nonreductive\textsubscript{2} physicalists or nonreductive\textsubscript{3} physicalists, and are not nonreductive\textsubscript{1} physicalists.

When Sydney Shoemaker says, for example, that "the view I favor is, then, a radically 'nonreductive' version of materialism" (1994: 56), it is clear from surrounding passages that he simply means to deny reductionism\textsubscript{2}, and does not mean to endorse the combination of physicalism on the one hand, and the denial of reductionism\textsubscript{1} on the other. In short, there is certainly nothing objectionable in Shoemaker's nonreductive physicalism: it is simply nonreductive\textsubscript{2} physicalism, and not nonreductive\textsubscript{1} physicalism.

Similarly, when Jerry Fodor, in a classic early paper, defends a view that has been widely interpreted as nonreductive physicalism, it is reductionism\textsubscript{3}, and not reductionism\textsubscript{1}, that is at issue. Fodor says for example, that "if psychology is reducible to neurology, then for every psychological kind predicate there is a coextensive neurological kind predicate, and the generalization which states this coextension is a law" (1974: 135). And he goes
on to say that "it is obvious to the point of self certification" that this kind of reductionism -- which is clearly reductionism₃ -- does not obtain. In other words, Fodor seems to be defending nonreductive₃ physicalism, and not nonreductive₁ physicalism.¹³

However, to conclude that nonreductive physicalism is in general consistent merely on the ground that the positions of Shoemaker and Fodor are consistent would be, I think, seriously mistaken. For there are nonreductive physicalists and there are nonreductive physicalists. In particular, while Shoemaker and Fodor clearly do not have reductionism₁ in mind, there are definitely philosophers who do.

The classic example of such a philosopher, I think, is Donald Davidson, at least the Davidson of 'Mental Events' (1980). This paper is of course difficult to interpret; however, I think that it is very plausible to view Davidson as endorsing the idea that physicalism is consistent with the denial of reductionism₁. In particular, the passages in which Davidson debunks reductive programs in general -- the passages in which he speaks of "the catalogue of philosophy's defeats" (1980: 217) -- suggest very strongly that, according to him, it is quite impossible to identify a mental property either with a physical or with a topic neutral property. But that rules out reductionism₁;

¹³ Our distinction between the three kinds of reductionism clears up a puzzle that many have noticed in Fodor's philosophy. In work from the 1970s, Fodor seemed to reject reductionism whole heartedly, but in work from the dark 1980s, he endorses reductionism. The answer is simply that, in the 1970s, by 'reductionism' Fodor tended to mean reductionism₁, whereas in the 1980s, by 'reductionism' he tended to mean reductionism₃. To illustrate, consider the discussion of what he calls a 'physicalistic theory of content'. The motivation for such a theory, says Fodor, is reductionism: "it is hard to see...how one can be a Realist about intentionality without being, to some extent or other, a Reductionist (1987: 97). But when he turns to what reductionism in this context amounts to, it is clear that Fodor has reductionism₁ in mind: A physicalist theory of content is a theory which "articulates in nonsemantic and nonintentional terms, sufficient conditions for one bit of the world to be about another bit" (1987: 98).
moreover, it suggests that, for Davidson, nonreductive physicalism is at least nonreductive\textsubscript{1} physicalism.\textsuperscript{14}

In sum then, on the basis of our tripartite distinction between reductionist theses, one might distinguish three versions of nonreductive physicalism: nonreductive\textsubscript{1} physicalism, nonreductive\textsubscript{2} physicalism and nonreductive\textsubscript{3} physicalism. Many nonreductive physicalists intend to endorse only the second or third of these positions, and it seems obvious enough that such positions are consistent. On the other hand however, such positions also seem rather uninteresting: after all, to hold them all you need to do is hold the topic neutral account of mental properties, and it is widely known that the topic neutral account is consistent with physicalism. What is interesting however, and what is not obvious, is whether philosophers such as Davidson are right to suppose that one can on the one hand be a physicalist and on the other suppose that mental properties are identical neither with physical properties nor with topic neutral properties. And it is this question, I suggest, that is the real question of nonreductive physicalism.

VII

We have seen, then, that the issue about nonreductive physicalism is properly understood as the issue of whether physicalism is consistent with the denial of reductionism, where that is taken to be the thesis that mental properties are identical either with physical or with topic neutral properties. Moreover, we

\textsuperscript{14} I do not mean to suggest here of course, that reductionism\textsubscript{1} is the only thesis of reductionism that interests Davidson. In particular, I do not mean to deny what is obviously true, namely that Davidson is interested in denying the possibility of psychophysical laws, i.e., that he is interested in denying the possibility of reductionism\textsubscript{3}. The point rather is that, when he talks of reductionism, Davidson casts his net very widely -- so widely, in fact, that he includes both reductionism\textsubscript{1} and reductionism\textsubscript{3}. For another example of Davidsonian nonreductive physicalism, see Schiffer 1992.
have seen that what gives this issue significance is that philosophers such as Davidson, as least as I read them, suppose that physicalism is consistent with the denial of reductionism so interpreted. What we have not seen, however, is whether Davidson is correct to suppose this; what we have not seen, in other words, is whether nonreductive physicalism is possible. What I should like to do in the remainder of the paper, therefore, is consider two lines of argument to the effect that Davidson and others are not correct, and thus that nonreductive physicalism is impossible.\textsuperscript{15}

The first is the argument of Kim's alluded to earlier in our discussion of necessary co-instantiation and strong supervenience. Kim's argument has three premises. The first is that physicalism is best formulated as a strong supervenience thesis. The second is that the strong supervenience of mental properties on physical properties entails that mental properties are necessarily co-extensive with physical properties. The third premise is that necessary co-extension is sufficient for property identity and hence for reduction. The conclusion of the argument is that physicalism entails reductionism.

Different philosophers have reacted in different ways to Kim's argument. Some argue that Kim is not entitled to the second premise: as we saw earlier, strong supervenience entails the necessary co-extension of mental properties only given certain closure assumptions. It therefore might be possible to reject these closure assumptions, and so avoid the argument. Others attack the third premise, and argue that necessary co-extension of properties is not sufficient for reductionism.\textsuperscript{16}

\textsuperscript{15} From here on, by 'reductionism' I will exclusively mean reductionism\textsubscript{1}, and will therefore eliminate the subscripts.

\textsuperscript{16} See, for example, Horgan 1994.
There are a number of problems with both these proposals, but in any case they are irrelevant. The main problem with Kim's argument is not in the second and third premises; the main problem is right at the beginning. As we saw previously, physicalism is not best formulated as a strong supervenience claim, it is best formulated as a weak_2 supervenience claim. Moreover, weak_2 supervenience, at least if it is suitably interpreted, does not entail the necessary co-extension of mental and physical properties. In other words, the first premise of Kim's argument is false, and the argument collapses.

VIII

Even if Kim's argument does not succeed in establishing that physicalism entails reductionism however, I nevertheless think that there is a second argument for this conclusion which is worthy of consideration. The first premise of this argument invokes the Humean principle that there are no necessary connections between distinct existences. The second premise is that physicalism involves necessary connections between mental facts and physical facts. It follows as a preliminary conclusion that if physicalism is true, mental facts are not distinct from physical facts. The third premise of the argument is that if mental facts are not distinct from physical facts, then reductionism is true. The overall conclusion is that one cannot be a physicalist without being a reductionist.17

Perhaps the most controversial step in this argument is the first premise that relies on the Humean principle that there are no necessary connections between distinct existences. I think there is no question that this principle is

---

17 Frank Jackson (1994: 29) suggests invoking the Humean ban on necessary connections between distinct existences to counter a position which admits supervenience but denies physicalism. I want here to invoke the ban against a related position, viz., a position which admits supervenience and physicalism, but denies reductionism.
true in some sense or other. However, there is no possibility of my defending
the principle here, and I will simply take it for granted in what follows. What I
will do instead is consider two different responses to the argument. The first
response is that the second premise is false because physicalism need not
involve necessary connections between mental facts and physical facts. The
second response is that the third premise is false because mental facts might fail
to be distinct from physical facts and yet at the same time reductionism might
not be true. I will argue that neither of these responses succeeds in combating
the Humean argument, and that to this extent this argument does establish
what Kim's argument does not, viz., that physicalism is reductive physicalism.

Response #1: Necessary connections not necessary.

Discussions of supervenience and physicalism -- and in this regard our
discussion was no different -- typically presuppose that, whatever else
physicalism is, it is at least the claim that mental facts entail the physical facts.
And of course this suggests that the second premise of the Humean argument
is right, and that if physicalism is true, then there are necessary connections
between mental facts and physical facts. However, might it not be possible to
deny this premise and articulate a notion of physicalism which does not
involve necessary connections between the mental and the physical?

The answer to this question, I think, is no. The difficulty in supposing
that physicalism does not involve the entailment of mental facts by physical
facts is that any attempt to formulate a doctrine that does not involve such an
entailment yields theses that are substantially weaker than the doctrine of
physicalism. Historically, the most influential version of physicalism that
eschews the entailment of the mental facts by the physical facts appeals to what
Jerry Fodor (1974) calls *token physicalism*. According to token physicalism,
every dated, spatially located particular -- every dated, spatially located object, event, state or process -- is a physical particular, where by 'physical particular' I mean a particular that instantiates a physical property. If you defend physicalism and mean by this token physicalism, then there is no question of the physical facts about a person entailing the mental facts. For it is consistent with token physicalism the connections between mental and physical properties are purely contingent.

Token physicalism is in some respects an attractive doctrine. It conforms to our intuition that there are no (actual) disembodied minds. Moreover, it makes sense of the view that Fodor calls the generality of physics: "roughly, the view that all events which fall under the laws of any special science are physical events and hence fall under the laws of physics" (1974: 127). Despite this, however, it is very hard to count token physicalism as a genuine version of physicalism. For token physicalism is a very weak claim. In particular, token physicalism is consistent with property dualism. That is, that every particular is a physical particular does not rule out the possibility that some particulars have irreducibly mental properties. But property dualism cannot be counted a form of physicalism. And thus, if token physicalism is consistent with non-physicalism, it is difficult to regard it as a physicalism.

The problem here is not simply intuition mongering about what counts as a 'genuine' physicalist claim. Every debate in philosophy of mind which concerns the truth of physicalism quite clearly proceeds on the assumption that property dualism is not a form of physicalism. The discussions of Frank Jackson's knowledge argument over Mary, for example, concern the question of whether "certain properties of certain mental states, namely those I've called qualia, are such that their possession or absence makes no difference to the physical world" (Jackson 1982: 473). Similarly, the debates concerning
intentionality concern the question of whether intentional properties can be reduced to other kinds of properties. In short, while token physicalism certainly rules out that there are irreducibly mental individuals, it does not rule out that there are irreducibly mental properties, and this means it is no form of physicalism. And this suggests that one cannot avoid the first premise of the Humean argument by adopting token physicalism.

Response #2: Not Distinct and Yet Not Reduced?

Let us suppose that the first two premises of the Humean argument are granted -- that is, let us suppose that physicalism requires necessary connections between mental facts and physical facts, and that there are no necessary connections between distinct existences. If follows of course that if physicalism is true, mental facts are not distinct from physical facts. But it does not follow that physicalism entails reductionism. In order to derive that conclusion, one needs the third premise of the Humean argument, namely that if mental facts are not distinct from physical facts, then reductionism is true. However, might this third premise not be denied? That is, might it not be suggested on the one hand, that mental facts are not distinct from physical facts, but on the other hand, that reductionism is false?

The main problem with this suggestion is simply that it is extremely hard to see how mental facts could fail to be distinct from physical facts without reductionism being true. In particular, if we consider the cases in which intuitively one fact F fails to be distinct from another fact F', it would seem that,

---

18 This shows the mistake in Colin McGinn's (1978) influential reply to Kripke's (1980) argument against physicalism. McGinn argues that token physicalism "can meet the requirements on an adequate rebuttal of the Cartesian challenge that Kripke lays down" (1978: 157). McGinn is certainly right that token physicalism avoid Kripke's arguments; but that isn't very interesting as a defence of physicalism since token physicalism is on its own too weak to be counted a variety of physicalism anyway.
in all of these cases, reductionism is true. In the simplest case, for example, a mental fact M fails to be distinct from a physical fact P because M is identical to P. In this case, however, it is clear that reductionism is true: if M is identical to P, the mental property which constitutes M is identical to the physical property that constitutes P. Or consider the more complicated case in which mental fact M fails to be distinct from physical fact P because M entails P but is not identical with it. Suppose for example that P is the fact of John's having physical property P* which has a certain causal role C, and M is the fact of John's being in pain. Then F entails M, and so is not distinct from it -- however this is so only because M is identical to the fact of John's having some property with causal role C. It is equally clear, then, that in this more complicated case, reductionism is true. More generally, if we take these intuitive cases to be our guide, then with failure of distinctness comes reductionism.

It might be objected, however, that our reliance of the intuitive cases is misleading. Our argument seems to be, as Stephen Schiffer has put it, that "it is illegitimate to invoke supervenience unless it can be shown to be a species of some familiar entailment" (1991: 186). However, as Schiffer goes no to say, surely this is too much to ask. In particular, surely it is too much to ask that the necessary relations inherent in supervenience be shown to be a species of some familiar entailment. Demanding this seems tantamount to demanding that physicalists come up with a proof or demonstration of the entailment of the mental by the physical. But this does not seem like a reasonable demand. After all, it is easy to imagine that we can provide no such a proof: for

19 I assume here what I take to be the standard account of fact-identity, viz., that the fact of a's being G is identical to the fact of b's being G iff a is identical to b and being F is identical to being G.
example, perhaps we simply do not know the proof; indeed perhaps we do not know it as a consequence of our natures, i.e., as a consequence of the fact that we are epistemically bound. On the other hand, however, that physicalists do not know such a proof, and therefore cannot provide one, is surely no objection to the idea that mental properties supervene on physical properties.

However, this reply misconstrues the point of the Humean argument. The point of the Humean argument is not to demand that a proof or demonstration of supervenience be produced by physicalists. The point rather is two-fold: first, to demand that if physicalism is true then mental facts cannot be distinct from physical facts, and, second, to argue that the only kinds of lack of distinctness we can imagine are cases in which reductionism is true. However, it is quite consistent with both of these points that physicalists cannot provide a proof of the entailment relations between mental and physical. In other words, then, one might perfectly reasonably hold to the view that providing a proof of supervenience is beyond what physicalists can manage; but this does not constitute a reply to the Humean argument.

Could the nonreductive physicalist nevertheless insist that failure of distinctness between mental facts indicates the existence of some unknown logical relation between mental and physical, which holds when physicalism is true, but can also hold in the absence of reductionism? Such a claim should be looked on with suspicion. It is of course reasonable in certain cases to suppose that we are ignorant of certain matters. Human limitation being what it is, we are often ignorant both of contingent facts and of necessary facts. But the ignorance the nonreductive physicalist is asking us to countenance here is of a different kind. It is not ignorance about whether one property is necessary related to another; reductionists should be happy to accept that we are sometimes ignorant about such things. And nor is it ignorance about particular
proofs or demonstrations for one property's being necessarily related to another; reductionists should be happy to accept that we might be ignorant about such proofs: after all, it is quite obvious that we can be ignorant about theoretical identification -- but theoretical identifications are ipso facto reductions. Rather we are being asked to accept that a logical relation of which we are ignorant -- a logical relation involved in supervenience -- is such that it is consistent with the falsity of reductionism. But I cannot see that this is a tenable position. Why should we accept that a relation of which we are ignorant must take a certain form? It is difficult to say in one tone of voice that we are ignorant of something, and in another tone of voice to say authoritatively that the thing of which we are ignorant must have such and such a property. 20

I should perhaps add that this last argument is not, and nor is intended to be, a knockdown argument against the suggestion that we might have supervenience without reduction. The considerations favoring the third premise of the Humean argument -- indeed the considerations in favoring the Humean argument itself -- are considerations of plausibility, not considerations which have the effect of ruling out all possible alternatives. Nevertheless it is worth emphasis that in assessing the fortunes of nonreductive physicalism we are interested not only in questions of reduction and non-reduction, but also in the broader questions of what the proper conception of mental and physical properties is, and of the why it is important that

20 It is important to emphasize here that our argument against nonreductive physicalism cross-cuts another issue that divides physicalists, viz., the issue of whether the necessary relations between mental and physical are known a posteriori or a priori. I have argued that supervenience entails reductionism, but I have not argued -- and nor does it follow from what I have argued -- either that the necessary relations involved in supervenience are known a priori, or that they are not. My intention here is to remain officially neutral on this much debated question. For arguments that any reasonable physicalism will inevitably assume an a priori or conceptual form, see Jackson 1980a, 1994, White 1986, and Botterell 1995. For the opposite view, see Byrne 1993, and Loar 1990.
physicalism is true. If physicalism can only be defended by adopting nonreductive physicalism and if nonreductive physicalism is committed to the kind of epistemological claims that we have just been considering then I think we need to reconsider the assumptions of our debate.

I conclude, then, that one cannot be a nonreductive physicalist in the interesting sense that one might hold physicalism together with the thesis that mental properties are identical neither with physical nor with topic neutral properties. One can of course hold physicalism at the same time as denying various other theses of reductionism; however, as we have seen, the resulting positions will not be very interesting since they will be indistinguishable from versions of reductive physicalism.

XI

At the beginning of the paper, I distinguished two kinds of dispute that arise for those who think that physicalism is to be stated as a supervenience thesis. The first dispute concerned the kind of supervenience thesis physicalists had in mind when they claimed that mental properties supervened on physical properties. The second dispute concerned the question of whether, in committing themselves to that supervenience thesis, physicalists are thereby committing themselves to reductionism.

My discussion of these related disputes has yielded two main conclusions. The first conclusion is that in defending the supervenience of mental properties on physical properties, physicalists are at least committed to a claim which on the one hand guarantees that mental facts are entailed by physical facts, and on the other hand, does not guarantee the reverse entailment. I suggested that weak\textsubscript{2} supervenience might be the kind of thesis that fits the bill.
The second conclusion is that in defending that kind of supervenience thesis, physicalists are committed to reductionism, so long as by ‘reductionism’ you mean the thesis that mental properties are identical either to physical properties or to topic neutral properties. 21

21 I am very much indebted to Andrew Botterell, Ned Block, Alex Byrne, Leonard Clapp, Robert Stalnaker, Judith Jarvis Thomson, and Zoltan Szabo for their comments on previous drafts of this paper.
Chapter Two
Nominalist Physicalism

I

In one formulation, the problem of intentionality is presented as concerning a particular class of properties, intentional properties. Intentional properties are those properties expressed by predicates formed from verbs of propositional attitude. For example, the verb 'believe' can form the predicate 'believes that snow is white', and this predicate expresses the intentional property, believing that snow is white. Similarly, the verb 'desire' can form the predicate 'desires that grass be green', and this predicate expresses the intentional property, desiring that grass be green. Given this characterization of intentional properties, the problem of intentionality is then the problem of explaining how it is possible that physical objects -- human beings, for example -- can have intentional properties. Thus, if we were to consider particular cases -- such as John's believing that snow is white, or Mary's desiring that grass be green -- the problem is to explain how it is possible that physical objects, like John or Mary, can have the property of believing that snow is white, or the property of desiring that grass be green.¹

The first thing to notice about the problem in this formulation is perhaps that it is not so very clear that it is a problem. For the question 'How is it possible that physical objects can have intentional properties?' would not in general be raised unless it were assumed that there was some prima facie reason for thinking that something could not be both a physical object and have

¹ An example of a philosopher who understands the problem of intentionality to be a problem of this form is Jerry Fodor. He remarks "What's wanted, -- for the geological properties as well as the psychological properties -- is just that we be able to understand how purely physical things can have them". See Fodor 1991 p181.
intentional properties. But it is not obvious that this assumption is true. After all, what exactly is the reason for thinking that a physical object cannot have an intentional property, and why is this reason so compelling that it gives rise to the problem of intentionality?

Philosophers suspicious of philosophical problems are inclined to take this kind of consideration very seriously, and as a consequence are inclined to think that there is no problem of intentionality. One such suspicious philosopher is Stephen Schiffer, and in his book *Remnants of Meaning* and in subsequent papers, he argues that "the questions which now define the philosophy of language (and I use that rubric broadly to include questions about intentionality) have false presuppositions" (1987: 271). In particular, Schiffer argues, there really is no difficulty in a physical object's having an intentional property, and thus we have no need to face the problem of intentionality.

What is interesting about Schiffer's argument that there is no problem of intentionality is that it does not come from the expected direction. It is natural to expect that someone who wants to say that there is no problem of intentionality would discuss what it is to be intentional and what it is to be physical and argue that on a proper understanding of these notions, there is no difficulty in a physical object's having an intentional property. Schiffer's idea, however, is that whether or not there is a problem of intentionality depends, not on the proper understanding of what it is to be intentional or physical, but rather on what positions are the correct ones in the metaphysics of properties. Schiffer argues that there are two conceptions of what properties are -- and

---


3 For an example of this kind of argument, see Tim Crane and D.H.Mellor 1991.
hence two conceptions of what intentional properties are -- and that only on what Schiffer thinks of as the incorrect conception is there a difficulty in a physical object's having an intentional property. If one adopts what is for Schiffer the correct conception of properties, then there is no difficulty, and hence, no problem.

Schiffer's strategy for avoiding the problem of intentionality is interesting for a second reason as well. For while it is not obvious that the problem of intentionality is a problem, it has normally been assumed that, for at least one group of philosophers, it is, viz., physicalists. Physicalists hold that intentional properties supervene on physical properties in the sense -- roughly -- that for every intentional property there is some physical property such that the instantiation of the latter is metaphysically sufficient for the instantiation of the former. Given that it is not obvious that intentional properties stand to physical properties in the necessary relations that supervenience requires -- and indeed, that there are substantial reasons to doubt that intentional properties stand in such relations at all4 -- it seems that physicalists might well raise the question 'How is it possible that physical objects have intentional properties?'. Schiffer himself is a physicalist, but he thinks that physicalists have no more reason to face the problem of intentionality than any other kind of philosopher. If physicalists attend to the metaphysics of properties in the way that he recommends, then, Schiffer thinks, for physicalists as for everybody else, there will be no problem of intentionality.

I will argue, however, that Schiffer's proposal that one might avoid the problem of intentionality if one adopts a particular conception of properties

4 According to Fodor, the "deepest motivation" for the problem of intentionality comes "from a certain ontological intuition: that there is no place for intentional categories in a physicalistic view of the world". See Fodor 1987 p97.
fails. I will not be concerned to argue that there is a problem of intentionality (or that there isn't). I will simply be concerned to show that Schiffer's remarks about the metaphysics of properties do not establish that there isn't a problem of intentionality, and, in fact, have no impact on that problem at all.

II

The first difference between Schiffer's two conceptions of properties emerges when we consider their contrasting accounts of certain claims about the existence and instantiation of properties. To adopt one of Schiffer's examples, according to the first conception of properties -- which he calls the non-pleonastic conception -- to say that there is the intentional property of believing that flounders snore is to be committed to the existence of certain kind of entity, something which is "objective, abstract and eternal" (1987: 144), and something which bears only a contingent relation to the predicate 'believes that flounders snore'. As Schiffer puts it, according to the non-pleonastic conception, the property of believing that flounders snore is "as ontologically and conceptually distinct from the predicate 'believes that flounders snore' as Saul Kripke is from the name 'Saul Kripke'" (1987: 144). Moreover, according to the non-pleonastic conception, to say that someone -- Ralph, in Schiffer's example -- has the property of believing that flounders snore is to say that Ralph stands in a particular relation to this objective abstract and eternal property.

In contrast, according to the other conception of properties that Schiffer considers -- which he calls the pleonastic conception -- to say that there is the

---

5 I ignore here the problem of saying what, on the non-pleonastic conception, properties are, and what, on the non-pleonastic conception, the instantiation of a property by an object comes to. The important points for our purposes will simply be that, on the non-pleonastic conception, properties are distinct from predicates, and the instantiation relation is a distinct relation from the relation of being true of. For discussion of these issues see David Armstrong 1978 Vol. I, as well as David Lewis 1983.
property of believing that flounders snore is not to say that there is something objective, abstract, and eternal; it is only to say that there is the \textit{predicate} 'believes that flounders snore'. Moreover, to say that Ralph has the property of believing that flounders snore is not to assert a relation between Ralph and an abstract entity; it is merely to say that the predicate 'believes that flounders snore' is true of Ralph.

The first difference between the two conceptions of properties that Schiffer considers, then, concerns the way in which we speak of properties: according to the non-pleonastic conception, to speak of properties is to speak of abstract entities; according to the pleonastic conception, to speak of properties is to speak only of predicates. The second difference emerges when we consider the truth conditions of the sentence:

(1) Ralph believes that flounders snore.

According to the non-pleonastic conception, the truth of this sentence inevitably involves the existence of the objective abstract and eternal property of believing that flounders snore. For according to this conception, (1) expresses a relation between the thing denoted by the name 'Ralph', namely Ralph, and the thing expressed by the predicate 'believes that flounders snore', namely, the objective, abstract and eternal property of believing that flounders snore. In particular, (1) is true if and only if Ralph instantiates this objective, abstract and eternal property.

In contrast, according to pleonastic conception, there is no abstract property of believing that flounders snore. \textit{A fortiori}, Ralph could not have this property. Do we then conclude that, on the pleonastic conception, (1) is false or is neither true nor false? No. According to the pleonastic conception, the truth conditions of sentences such as (1) do \textit{not} involve commitment of properties in the non-pleonastic conception. Rather, (1) is true if and only if the
predicate 'believes that flounders snore' is true of Ralph. From the point of view of the pleonastic conception, the predicate 'believes that flounders snore' can be true of Ralph even if there is no abstract property of believing that flounders snore, and thus (1) can be true in the absence of that property.6

The second difference between the two conceptions of properties, then, emerges when we consider their contrasting accounts of the truth conditions of ordinary sentences such as (1). The pleonastic conception of properties asserts not only that to speak of properties is to speak merely of predicates; it also asserts that sentences such as 'Ralph believes that flounders snore' are true just in case 'believes that flounders snore' are true of Ralph. On the other hand, the non-pleonastic conception asserts not only that to speak of properties is to speak of abstract entities; it also asserts that the truth of sentences such as

6 Schiffer's pleonastic conception of properties as I have presented it is most similar to the version of nominalism which Armstrong calls 'predicate nominalism'. See Armstrong 1978, p13. I do not mean to suggest, however, that there are not other ways to interpret what Schiffer says concerning the metaphysics of properties. Indeed, I do not think that Schiffer himself is entirely clear on which of a number of possible views he intends. Predicate nominalism seems plausibly attributed to Schiffer in those parts of Remnants of Meaning that deal with the questions of physicalism and reduction that I want to focus on. In other parts of Remnants, however, Schiffer suggests somewhat tentatively that one might appeal to substitutional quantification to provide an analysis of (1) that avoids commitment to properties; such a view would apparently differ from the one in the text in that it would not paraphrase (1) into (2). (This substitutional quantification suggestion is repudiated in later papers. See Schiffer 1994.) A third possibility is to understand the pleonastic conception of properties as analogous to contemporary deflationary theories of truth, properties and facts. However, as Paul Horwich pointed out to me, these latter theories tend to be weaker than the pleonastic conception of properties since they do not attempt to analyze talk of properties into talk of predicates. For our purposes, however, it is sufficient to focus exclusively on the version of nominalism discussed in the text. First, predicate nominalism has the advantage of being a rather clear and frank version of nominalism. Second, and more important, the existence of other versions nominalism (or deflationism) does not, I think, affect the main point I wish to make, which is that nominalism or deflationism in general cannot make the problem of intentionality disappear. For a recent defense of the deflationary theory of truth, see, Horwich 1990; for deflationary theories of facts and properties, see Horwich 1993, as well as Schiffer, 1987, ch. 6; for a discussion of the relation between the deflationary theory of truth and the problem of intentionality, see Boghossian 1990, and Devitt 1990; for a discussion of the relation between the deflationary theory of truth and the problem of linguistic meaning, see Johnston 1988.
'Ralph believes that flounders snore' inevitably involves the existence of such entities.

We can summarize the difference between the pleonastic and the non-pleonastic conceptions by considering their differing answers to the question 'What makes it the case that 'believes that flounders snore' is true of Ralph?' According to the non-pleonastic conception, what makes it the case that 'believes that flounders snore' is true of Ralph is (a) that there is the property of believing that flounders snore; (b) that this property is expressed by 'believes that flounders snore'; and (c) that Ralph instantiates this property. By contrast, according to the pleonastic conception, there is no property of believing that flounders snore which makes it the case that 'believes that flounders snore' is true of Ralph. Rather, this predicate simply is true of Ralph, and that is all there is to be said on the matter.7

Now Schiffer argues that on neither conception of intentional properties is there a problem of intentionality. I suggest that we can best state his argument if we first consider a thesis which explicitly connects the idea that physical objects have intentional properties with there being a problem of intentionality:

(I) There is a problem of intentionality if and only if physical objects can have intentional properties.

7The example of Ralph concerns an intentional property, the property of believing that flounders snore. And it is properties of this kind that are important in connection with the problem of intentionality. However, it should not be thought that either the pleonastic or the non-pleonastic conception of properties is a view about intentional properties alone and is not a view about non-intentional properties. Rather, Schiffer intends both conceptions to be quite general accounts of the existence of properties, and of what it means to say that an object has a certain property. Thus, to adopt another of Schiffer's examples, what goes for there being the property of believing that flounders snore goes also for there being the property of being humble, and what goes for 'Ralph believes that flounders snore' also goes for 'Mother Theresa is humble'.

Given Schiffer's distinction between the two conceptions of properties, it follows that (I) admits of two interpretations, depending on whether the claim that physical objects have intentional properties is to be interpreted according to the non-pleonastic conception of properties, or according to the pleonastic conception. Schiffer thinks that (I) is true under the non-pleonastic interpretation, and false under the pleonastic interpretation. But, he thinks, in either case one can derive the conclusion that there is no problem of intentionality.

Let us first suppose that the claim that physical objects can have intentional properties is to be interpreted according to the non-pleonastic conception of properties -- we might put this by saying that physical objects can non-pleonastically have intentional properties. Then we have the following interpretation of (I), which we can call (I-np):

(I-np) There is a problem of intentionality if and only if physical objects can non-pleonastically have intentional properties.

With the help of (I-np), the first part of Schiffer's argument to the claim that there is no problem of intentionality can be stated very simply. For Schiffer thinks that:

(I-np) is true

But he also thinks that the right hand side of (I-np) is false. For Schiffer agrees with the pleonastic conception that there are no properties at all;\(^8\) it follows, of course, that there are no intentional properties either, and thus it follows that:

It is not the case that physical objects can non-pleonastically have intentional properties.\(^9\)

---

\(^8\) For example: "I do not recognize any nonpleonastic properties....I favour nominalism across the board, and especially do not think that properties are needed as the semantic values of predicates" (1987:145). For further discussion, see 1987: 234-5.
But of course, if (I-np) is true, and if physical objects cannot non-pleonastically have intentional properties, it follows that there is no problem of intentionality.

Alternatively, let us suppose that the claim that physical objects can have intentional properties is interpreted according to the pleonastic conception of properties -- we might put this by saying that physical objects can merely pleonastically have intentional properties. Then we have an interpretation of (I) which we might call (I-p):

\[(I-p) \quad \text{There is a problem of intentionality if and only if physical objects can merely pleonastically have intentional properties.} \]

With the help of (I-p), the second part of Schiffer's argument can also be stated simply. For, as I understand him, Schiffer holds that:

\[(I-p) \quad \text{is false.} \]

For Schiffer rejects the non-pleonastic conception of properties, but endorses the pleonastic conception. Moreover, on that conception, he agrees that physical objects have intentional properties. He endorses, then, the claim that:

\[\text{Physical objects can merely pleonastically have intentional properties.} \]

But of course, if (I-p) is false, physical objects can merely pleonastically have intentional properties, it again follows that there is no problem of intentionality.

In sum, Schiffer's is a two part argument. It begins with the idea that there are two conceptions of properties, the pleonastic and the non-pleonastic.

---

9 I assume here that the claim that there are no properties in the non-pleonastic conception is, if true, necessarily true. If there are no intentional properties, then, it follows that physical objects cannot have such properties.

10 I am assuming here that it is part of the pleonastic conception of properties that if an object pleonastically has properties then it is not the case that the object non-pleonastically has properties -- hence the 'merely' in the statement of (I-p).
On the non-pleonastic conception of properties, (I-np) is true, but physical objects do not have intentional properties; it follows that there is no problem of intentionality. On the pleonastic conception of properties, (I-p) is false, but physical objects do have intentional properties; it again follows that there is no problem of intentionality.

Now there are clearly a number of different ways that one could respond to this argument. One might, for example, criticise the distinction Schiffer draws between the pleonastic and non-pleonastic conceptions of properties. If there were no such distinction, or if the distinction were less clear than Schiffer supposes, then his argument would collapse, for it is crucial to his argument that there are two clearly different interpretations of (I). Alternatively, one might argue that Schiffer is mistaken to be a nominalist, to deny that there are any properties in the non-pleonastic conception. If nominalism in general were mistaken, it would become less obvious why one should deny that physical objects non-pleonastically have intentional properties. But if physical objects non-pleonastically had intentional properties, then Schiffer's argument collapses, for he concedes that if they did then there would be a problem of intentionality; it is just that he thinks they don't, so there isn't.

But I want to pursue a different strategy. Instead of concentrating on Schiffer's distinction between the pleonastic and non-pleonastic conceptions of properties, or on the correctness of nominalism, I want to take up the following question: if we suppose that there is problem of intentionality on the condition that physical objects non-pleonastically have intentional properites, is there any good reason to suppose that there is not a problem of intentionality on the condition that physical objects merely pleonastically have intentional

---

11 For examples of this line of criticism, at least as applied to intentional properties, see Fodor 1991 and Kim Sterelny's (1989) review of Remnants of Meaning.
properites? I will argue that there is no good reason to suppose this, and thus, contrary to Schiffer's argument, the question of whether physical objects non-pleonastically have intentional properties, or whether they merely pleonastically have intentional properties, is irrelevant to the question of whether there is a problem of intentionality.

III
Let me begin by considering thesis (I) and its two interpretations, (I-np) and (I-p). It is crucial to Schiffer's argument that (I-np) is true while (I-p) is false. But the first thing to be said in response to this is that, prima facie anyway, whatever reason there is to be invoked in support of (I-np), there is a parallel reason to be invoked in support of (I-p).

After all, what would be the reason for thinking that (I-np) is true? One would think that (I-np) is true, I suggest, if one thinks that there is some difficulty in the idea of a physical object's non-pleonastically having an intentional property, a difficulty which would give rise to the problem of intentionality. If there were such a difficulty, then, from the assumption that physical objects did non-pleonastically have intentional properties, it could be concluded that there is a problem of intentionality. 12

Let us suppose then that there is some difficulty in the idea of a physical object's non-pleonastically having an intentional property. Prima facie, it is very hard to see why there should not also be a difficulty in the idea of a physical object's merely pleonastically having an intentional property. After all, to say that physical objects merely pleonastically have intentional properties

12 Strictly speaking, this kind of consideration provides support only for one half of (I-np), the half that says that the fact that physical objects have intentional properties entails the fact that there is a problem of intentionality. This complication does not matter, I think, to the point I want to make.
is to say that intentional predicates are true of physical objects. But why should the shift from properties to predicates make any difference? On the assumption that there is a difficulty in the idea of Ralph's non-pleonastically having the property of believing that flounders snore it is natural to think that there should also be a difficulty in the idea of the predicate 'believes that flounder snore' being true of Ralph. In fact, it is reasonable to suppose that whatever it is that makes trouble for the one will make trouble for the other. In short, then, and contra Schiffer's argument, if we suppose that (I-np) is true, there seems every reason to suppose that (I-p) is true as well.

It might be replied that this depends on what the difficulty is in the idea of a physical object's non-pleonastically having an intentional property. Presumably some difficulties will survive the transition from non-pleonastic to pleonastic and others won't. This is of course true; but I do not think it affects the fact that there is no prima facie reason to think that Schiffer's argument will be successful. For it is important to point out that the difficulty that Schiffer himself finds in the idea of a physical object's having an intentional property seems precisely the sort of difficulty that remains unaffected by the distinction between the pleonastic and non-pleonastic conceptions of properties.

For Schiffer, the difficulty has its source in the possibility that intentional properties are irreducible, that is, in the possibility that there are no physical properties to which intentional properties are reducible. On the one hand, Schiffer thinks, it is hard to see how intentional properties can be reduced to physical properties. On the other hand, he thinks, it is also hard to see how physical objects can have intentional properties if those properties are not so reducible to physical properties. The upshot is that there is a difficulty in seeing how physical objects can have intentional properties at all. This is the difficulty which leads Schiffer to think that if physical objects non-
pleonastically had intentional properties, then there would be a problem of intentionality.

Now Schiffer of course might be right or wrong to find difficulty in the idea of a physical object's having an irreducible mental property. Regardless of whether he is right or wrong however, it difficult to see why the idea that irreducible mental predicates are true of physical objects should be any less troublesome than the idea that physical objects have irreducible mental properties. And this suggests again that, prima facie, there is no reason to suppose that (I-np) is true while at the same time supposing that (I-p) is false.

IV

So far, then, I have suggested that while we might grant to Schiffer that there are two interpretations of (I), there appears no prima facie reason to grant to him that (I) is true under one of these interpretations -- the interpretation expressed by (I-np) -- but false under the other -- the interpretation expressed by (I-p). Schiffer thinks that (I-np) is true because he thinks there is a difficulty in a physical object's having an irreducible intentional property. But if there is such a difficulty, there would prima facie be a parallel difficulty in an irreducible intentional predicate being true of a physical object.

I want now to examine in more detail the question of reductionism. We have seen that Schiffer thinks that there is difficulty in the idea of a physical object's having an irreducible mental property, but that there is no parallel difficulty in the idea of an irreducible mental predicate being true of a physical object. However, why does Schiffer think this? What, after all, is the problem that Schiffer finds in a physical object's having an irreducible mental property, and why is there no parallel problem in there being an irreducible mental predicate true of a physical object?
Schiffer's argument for reductionism appeals to considerations of mental causation. Schiffer invites us first to agree that sentences of the form:

(a) Ava stepped back because she believed that a car was coming.

are, if true at all, always true in conjunction with a sentence drawn from a different class of sentences, namely sentences of the form:

(b) Ava stepped back because she was in neural state N.

Schiffer then says that, regardless of one's metaphysical commitments concerning properties, one must explain how it is that, if a sentence such as (a) is true, there is always a sentence such as (b) which is true as well. We might put this more briefly by saying that, regardless of one's metaphysical commitments, one must explain the joint truth of (a) and (b).

Now, Schiffer argues that the only way for the friend of the non-pleonastic conception of properties to explain this fact -- the fact of the joint truth of (a) and (b) -- is to suppose that the property of believing that a car was coming is reducible to the property of being in neural state N. Schiffer draws the conclusion from this that the friend of the non-pleonastic conception of properties must agree that if a physical object has an intentional property, that property must be reducible to a physical property.

Now one can certainly imagine reasons to be sceptical about this line of argument. However, let us set such reasons aside. The important point for our purposes is not whether Schiffer's argument for reductionism is sound, but rather whether the same line of reasoning will go through if one is working with the pleonastic conception of properties. Schiffer thinks not. In particular, he thinks there are at least two explanations of the joint truth of (a) and (b) which do not amount to reductionism, which are available to a friend of the

---

13 For some of these reasons, see Antony 1991.
pleonastic conception of properties, but which were not available to the friend of the non-pleonastic conception.

The first of these -- which I will call the happy coincidence view -- is the view that "it is simply a happy coincidence that (a) and (b) are true together" (187: 173); that is, that it is just a coincidence that every time a sentence such as (a) is true, there is a sentence such as (b) which is true as well. Now this happy coincidence view is an extremely odd position. After all, can it really be mere good luck that we are in a world in which every time it is true that someone stepped back because she saw a car coming, it is also true that she stepped back because she was in a particular neural state? That kind of coincidence is surely altogether too happy to be a coincidence.

Schiffer agrees that the happy coincidence view is implausible for the friend of the non-pleonastic conception of properties, but he suggests the situation is different for the friend of the pleonastic conception. According to Schiffer, if we understand "the genesis of propositional attitudes and propositional attitude concepts" this will "remove...what threatens to be an ontological mystery" (1987: 170). In referring to the "genesis of propositional attitude concepts", Schiffer means, I think, to refer to what he takes to be what is most important about these concepts, namely that they enter into what he calls a "reliable predictive practice" (1991: 14). Amplifying on this idea, Schiffer says:

it is constitutive of our propositional attitude concepts that when we believe that a person has certain beliefs and desires, and certain other (not readily articulable) conditions obtain, then we form expectations about what the person will do and about the reason for which he or she will do it. (1987: 172)
On the basis of these expectations, we find ourselves "in a position to reliably predict someone's behaviour on the basis of propositional attitudes we ascribe to him" (1991: 14). In short, Schiffer seems to suggest that once it is acknowledged that our central purpose in ascribing propositional attitudes is to explain and predict the behaviour of others then the "ontological mystery" of the joint truth of the sentences (a) and (b) will vanish, and the way will be clear for the view that it is just a happy coincidence that these two sentences are true together.

But this suggestion is shot through with difficulties. First, it is completely unclear how the mere fact that we engage in a reliable predictive practice when we ascribe propositional attitudes to people has any bearing on the plausibility of the happy coincidence view. Schiffer is certainly correct that we engage in this practice. After all, this merely means that we have certain goals in asserting intentional sentences, and that we have a legitimate confidence in achieving those goals. However, whatever our goals are in asserting intentional sentences, those sentences nonetheless might be true. And if they are true, we will still want to know how their truth bears on the truth of various non-intentional sentences. For example, suppose it is my purpose to explain and predict Ava's behaviour when I say 'Ava believes there is a car coming'; suppose, that is, that I am engaging in a reliable predictive practice when I assert this sentence of Ava. Even so, the sentence may be true. And, if it is true, we will want to know how the truth of that sentence bears on the truth of various other sentences, such as, for example, 'Ava is in neural state N'. To suggest that these sentences are true together by coincidence does not seem like much of an answer. Second, even if Schiffer's point that we engage in a reliable predictive practice somehow bestowed credibility on the happy coincidence view, it is unclear what role the distinction he draws between the
two conceptions of properties plays in all this. After all, why can a friend of the non-pleonastic conception of properties not also appeal to the idea that our practice of attributing attitudes is a reliable predictive practice?

Furthermore, there is an independent problem with Schiffer's idea that the friend of the pleonastic conception of properties might appeal to the happy coincidence explanation of the joint truth of (a) and (b). As I noted at the outset, Schiffer promises, among other things, to show that physicalists have no need to face the problem of intentionality, that one can be a physicalist without worrying about the problem of intentionality. However, it is quite obvious that the happy coincidence view is not a form of physicalism. A minimal form of physicalism must respect the supervenience of the intentional on the physical. But the happy coincidence view violates supervenience. Suppose again that it is just a coincidence that 'believes that flounders snore' and 'is in such and such a neural state' are both true of Ralph. Then it follows that there is a possible world in which every physical sentence true in our world is true, but that 'believes that flounder snore' is true of nothing, and therefore not true of Ralph. This latter possibility is allowed by the happy coincidence view, but is inconsistent with physicalism. And again, there is no reason to think that this violation of physicalism is affected by a pleonastic interpretation or by a non-pleonastic interpretation of the properties involved.

There doesn't seem much, then, to the suggestion that the friend of the pleonastic conception can appeal to the happy coincidence view to avoid Schiffer's causal argument for reduction while the friend of the non-pleonastic conception cannot. First, the happy coincidence view is not a very good explanation of the joint truth of (a) and (b), so it is unclear that the happy coincidence view avoids the argument for reduction. Second -- and this point is the more important one for our purposes -- even if the happy coincidence
view did avoid the argument for reduction, the view is available to a friend of either conception of properties. In short, the happy coincidence view is not a particularly happy view.

However, Schiffer suggests that there is also a second explanation of the joint truth of

(a) Ava stepped back because she believed that there was a car coming
and
(b) Ava stepped back because she was in neural state N

which is unavailable to the friend of the non-pleonastic conception, but is available to a friend of the pleonastic conception. This is the view that the truth of (a) supervenes on the truth of (b) but is not reducible to it. Let us call this the supervenience without reduction view.14

Schiffer argues that the supervenience without reduction view is unavailable to the friend of the non-pleonastic conception. He remarks that, if one accepts the thesis there are properties, the supervenience without reduction view is "obscurantist in the extreme" (1987: 154), and compares the view to G.E. Moore's idea that moral properties supervene on natural properties without being reduced to them:

How could being told that non-natural moral properties stood in the supervenience relation to physical properties make them any more palatable? On the contrary, invoking a special primitive metaphysical relation of supervenience to explain how moral properties were related to physical properties was just to add mystery to mystery, to

14 The supervenience without reduction view of the joint truth of (a) and (b) is obviously incompatible with the happy coincidence view, so the question arises: Which view does Schiffer himself support? The answer is: both at different times. In Remnants of Meaning he supports the happy coincidence view (p173). In later papers, however, he supports supervenience. See, Schiffer 1991b p185.
cover one obscurantist move with another. I therefore find it more than a little ironic, and puzzling, that supervenience is nowadays being heralded as a way of making non-pleonastic, irreducibly non-natural intentional properties cohere with an acceptably naturalistic solution to the mind-body problem (1987: 153-4)

On the other hand, continues Schiffer, the supervenience without reduction view is available to the friend of the pleonastic conception of properties. In particular, he says (1987: 165), once one rejects the non-pleonastic conception of properties, one can endorse this "mild enough supervenience claim":

(S) Given that Ava believes that a car is coming toward her, she also believes this in every possible world that is physically indistinguishable from the actual world.

Given the truth of supervenience theses such as (S), one might easily explain how, whenever a sentence such as (a) is true a sentence such as (b) is true as well.

And thus, Schiffer concludes, (1987: 166) one can "answer the charge that an argument can be launched against [my view] that parallels the one against nonpleonastic irreducible belief properties".

I think there are serious reasons to doubt that supervenience claims of the form (S) -- which Jaegwon Kim calls global supervenience claims -- provide a reasonable answer to the problem of explaining the joint truth of (a) and (b).15 Once again, however, let us set such worries aside. The important point

15 The problem with (S) is as follows. (S) applies only to worlds that are physically indistinguishable from ours, and is completely silent about worlds that are physically distinguishable. However, there are worlds that are physically distinguishable from ours, but only in the most trivial ways. Let us suppose that, in our world, Wa, there is a pebble on the moon located in a certain position. Now consider another world, Wp, physically exactly like ours with but one difference: this pebble is located two inches from position it occupies in our world. Wp is physically distinguishable from our world. However, if Wp is physically distinguishable from ours, it is consistent with (S), that, in Wp, Ava does not believe that there is a car coming. Indeed, it is consistent with (S) that no-one has any beliefs at all! For this kind of criticism of global supervenience claims, see Kim 1993, p85
for our purposes is not whether (S) and its ilk provide a serious answer to the question of the joint truth of (a) and (b); the important point is to see whether Schiffer is right that (S) is not open to the very same charges that were levelled against supervenience claims that a friend of the non-pleonastic conception of properties might make.

But once again, it is very hard to see how Schiffer could be right that it is not. In the first place, it is quite mistaken to suppose that (S) is not available to the friend of the non-pleonastic conception of properties. Schiffer says that (S) "eschew[s]" properties (1987: 165), thereby suggesting that the friend of the non-pleonastic conception might not advance a supervenience claim such as (S). But in fact (S) does not eschew properties. On the contrary, (S) is neutral about the existence of properties, since it admits of both a pleonastic and a non-pleonastic reading. Moreover, there is no reason to think that the variety of supervenience claim that one adopts must be dictated by what metaphysical approach to properties one finds congenial. In short, if one likes global supervenience claims such as (S) one can have them -- even if one also likes the non-pleonastic conception of properties.

Second, there is no reason to suppose that supervenience claims such as (S) are immune to the criticisms Schiffer levels against other supervenience claims. Schiffer says that it piles mystery on mystery to suppose that there is a "primitive form of entailment" (1987:153) between physical properties and intentional properties of the kind required by supervenience without supposing that intentional properties are reducible to physical properties. But that kind of remark is perfectly applicable in the case of (S). For consider Ava. According to (S), if Ava is in a world physically indistinguishable from this one, then, as a matter of necessity, she believes that a car is coming. In other words, Ava's being in a world physically indistinguishable from this one entails that
she believes that there is a car coming. Schiffer wants to argue that, in spite of this, (S) does not imply that intentional properties pleonastically interpreted are reduced to physical properties pleonastically interpreted. But why is this proposal not to pile mystery on mystery when an exactly parallel proposal, according to Schiffer, is?

As far as Schiffer's argument for reductionism goes, then, the supervenience without reduction view is on all fours with the happy coincidence view. Both are available to friends of both conceptions of properties, and neither are attractive as accounts of the joint truth of (a) and (b). More generally, and to return to the main line of argument, Schiffer is mistaken to suppose that there is any less of a problem with irreducible mental predicates than there is with irreducible mental properties. If the Schiffer's argument for reductionism is sound, there is as much a problem with irreducible properties as there is with irreducible predicates.

VI

To this point, my concern has been with the theses (I-np) and (I-p). Schiffer thinks that (I-np) is true because he thinks that there is a difficulty with irreducible mental properties, a difficulty that arises from considerations of mental causation. On the other hand, Schiffer thinks that (I-p) is false because he thinks that there no parallel difficulty with irreducible mental predicates. As against this, however, we have seen that there is no reason to accept that (I-np) and (I-p) differ in truth value. To the extent that considerations of mental causation drive us to reductionism in non-pleonastic case, they drive us to reductionism in the pleonastic case; moreover, to the extent that difficulties about reductionism give rise the the problem of intentionality in one case, they give rise to that problem in the other.
I want now to turn to a different part of Schiffer’s argument for there not being a problem of intentionality: the premise that it is not the case that physical objects can non-pleonastically have intentional properties. One of Schiffer’s arguments for this premise is nominalism, and I have already sworn off any discussion of the correctness of nominalism. But Schiffer has a second argument for this premise, and this second argument is relevant to the question we have set for ourselves, namely, whether the distinction between pleonastic and non-pleonastic is relevant to the problem of intentionality.

The argument derives once again from Schiffer’s concerns about the irreducibility of intentional properties. To establish that intentional properties are reducible to physical properties, Schiffer thinks, it would be necessary to state in physical terms necessary and sufficient conditions for a physical object to have an intentional property; however, to demand that this be done is an "invitation to do the impossible" (1987:179). Schiffer summarizes this by saying that, if there are intentional properties, those properties are not reducible to physical properties.

On the other hand, the idea that physical objects have irreducibly intentional properties is, as Schiffer puts it, "not merely obscurantist but magical" (1987:11), and to suppose that physical objects have irreducibly intentional properties is to "renounce the scruples of the natural scientist", and, in Quine’s words which Schiffer quotes with approval, to "just surface listlessly to the Sargasso Sea of mentalism" (1987:142). Schiffer summarizes this point

16 Schiffer’s confidence that it is impossible to provide in physical terms necessary and sufficient conditions for a physical object to have a mental property derives in part from his arguments in the first five chapters of Remnants of Meaning, which criticise many current attempts to provide such conditions. Of course, even if Schiffer is correct that many -- or even all -- current attempts to provide such conditions fail, it scarcely follows that it is impossible to do so.

17 The quotation is from Quine 1975 p91.
by saying that, if there are intentional properties, those properties are reducible to physical properties.

These two points yield the following argument, which we can call the *property argument*:

(i) If there are intentional properties, those properties are not reducible to physical properties.

(ii) If there are intentional properties, those properties are reducible to physical properties.

Therefore,

(iii) There are no intentional properties.

If this argument is sound, we clearly have a very good reason to believe Schiffer's premise that it is not the case that physical objects non-pleonastically have intentional properties. For if there are no intentional properties, then it's not the case that physical objects non-pleonastically have such properties.

What is important for our purposes, however, is not the soundness of the *property argument* so much as whether it is possible to construct an analogous and equally sound argument consistently with the pleonastic conception of properties. This analogous argument would be exactly like the *property argument* except that the word `predicate' is substituted for the word `property' throughout:

(i*) If there are intentional predicates, those predicates are not reducible to physical predicates.

(ii*) If there are intentional predicates, those predicates are reducible to physical predicates.

Therefore,

(iii*) There are no intentional predicates.
Let us call this argument the predicate argument to distinguish it from the original property argument. Of course Schiffer thinks the property argument is sound. But then our question is the following: if we suppose that property argument is sound, is there any reason to suppose that predicate argument is unsound?

The question is very important for Schiffer's argument against there being a problem of intentionality. For let us suppose that both the property and the predicate argument were sound. Then we could derive the conclusion, not only that there are no intentional properties, but also that there are no intentional predicates. But it was no part of Schiffer's plan to deny that there are intentional predicates. Quite the contrary, Schiffer asserts that some intentional predicates are true of physical objects. For example, he thinks that 'believes that flounders snore' is -- or might be -- true of Ralph. Indeed, it was a presupposition of the argument I set out in section II that physical objects merely pleonastically have intentional predicates, that is, that intentional predicates are true of physical objects.

Schiffer thinks there is no reason to suppose that the predicate argument is sound in part because of the considerations from mental causation that we discussed in the previous section. Schiffer thinks these considerations support the second premise of the property argument, but do not similarly support the second premise of the predicate argument. As we have seen, however, there is no reason to suppose that Schiffer is right to think this.

But Schiffer has a second reason for thinking that the predicate argument is unsound while the property argument is sound. He thinks that the conclusion of the predicate argument -- the denial that there are intentional predicates -- is absurd while the conclusion of the property argument is not absurd: "What is the opposite number of the argument now that only belief
predicates are at issue? it could hardly be that there are no belief predicates!" (1987:166).

But it is difficult to see how Schiffer could be right about this. If by 'intentional predicate' -- or, in Schiffer's phrase, 'belief predicate' -- you mean sounds in the air and marks on a page, then of course one cannot deny that there are intentional predicates. However, one might very well deny of those sounds and marks that they are intentional predicates; one might very well deny, that is, that these sounds and marks are expressions literally true or false of physical objects. In that sense, one might very will deny that there are intentional predicates. Indeed, this is a denial with precedent. Gilbert Ryle adopts just such a position in The Concept of Mind when he says that sentences such as 'John Doe knows French' are "neither reports of observed or observable states of affairs, nor yet reports of unobserved or unobservable states of affairs". I take it that what Ryle meant by this is that expressions such as 'John Doe knows French' are not declarative sentences which are true or false. The same position might be expressed by saying that expressions like 'believes that flounders snore' are not predicates which are true or false of physical objects.

Now it might be replied that Ryle's non-factualist view that expressions like 'believes that flounders snore' are not predicates which are true or false of physical objects is, if perhaps not absurd, then extremely unattractive, and thus that if the predicate argument yields non-factualism as a conclusion, then that argument must be resisted. But the response to this is simply that the

---

18 See Ryle 1963 p120. For more recent discussion of a view like Ryle's, see Boghossian 1990, and Devitt 1991.

19 The label 'non-factualism' is due in this context to Boghossian, 1990. Likewise the label 'error theory'.
conclusion of the property argument is just as unattractive as Rylean non-factualism. Let us suppose that we accept that the property argument is sound, and therefore arrive at the view that there are no intentional properties. Then, if the non-pleoanastic conception of properties is right, we must conclude that all sentences of the form 'Ralph believes that flounders snore' are false, or at least are not true. For, according to the non-pleonastic conception, what makes it the case that a sentence such as 'Ralph believes that flounders snore' is true is that Ralph instantiates the property expressed by 'believes that flounder snore'. But of course if there is no such property, then Ralph cannot instantiate it, and the sentence is false. I take it that this kind of error theory is just as unattractive as the Rylean non-factualist view that there are no intentional predicates which are true or false of physical objects.

In sum, if the predicate argument is sound, it results in non-factualism; on the other hand, if the property argument is sound it results in the error theory. These are different positions; but I do not think that they are different enough to support Schiffer's contention that the conclusion of the predicate argument is absurd, while the conclusion of property argument is not. More generally, if we agree with Schiffer that the property argument is sound and conclude with him that it is not the case that physical object's non-pleonastically have intentional properties, we should equally think that the predicate argument is sound and and conclude that it is not the case that physical object's merely pleonastically have intentional properties. However, if we do that, then Schiffer's argument collapses.

VI

I began by distinguishing two halves of Schiffer's argument for the conclusion that there is no problem of intentionality. The first half was that the thesis (I) is
true under one interpretation -- the interpretation expressed by (I-np) -- but that it is not the case than physical objects non-pleonastically have intentional properties; the second half was that the thesis (I) is false under another interpretation -- the interpretation expressed by (I-p) -- but that it is the case that physical objects merely pleonastically have intentional properties. I have been arguing that there is no way to fit the two halves of this argument together. For one thing, whatever reason there may be for thinking that (I-np) is true, there is parallel reason for think that (I-p) is true too. For another, one of Schiffer's main arguments for the conclusion that it is not the case that physical objects non-pleonastically have intentional properties -- the property argument -- also yields the conclusion that it is not the case that physical objects merely pleonastically have intentional properties. Hence, if we believe one of Schiffer's premises on the basis of the property argument, we are obliged to give up another of his premises.

I want now to close the paper by considering a somewhat more general issue that is raised by our discussion. There is -- as we have noted -- a problem of intentionality only if there is a difficulty in the idea of a physical object's having an intentional property. I have been concentrating on the difficulty Schiffer perceives in the idea of physical object's having an intentional property -- the difficulty having to do with reductionism -- and have argued that the distinction between the pleonastic and non-pleonastic conception is irrelevant to this difficulty. This difficulty about intentional and physical properties remains, regardless of what conception of properties is invoked, and so the distinction between these conceptions of properties is irrelevant to the problem of intentionality.

One might wonder, however, whether there is another difficulty in the idea of a physical object's having a intentional property, a difficulty on which
Schiffer's distinction may be supposed to have an impact. And I think there is. Quite independently of the property argument, one might be disposed to reason as follows. It is just sordid semantics (it might be argued) to think that, corresponding to the meaningful expressions 'is red' and 'believes that flounders snore', there are the properties being red and believing that flounders snore. And it is just mysterious metaphysics (it might be continued) to suppose that these properties stand in the instantiation relation to physical objects. After all, what is the instantiation relation? Is it a two-place property, corresponding to the English word 'have'? If it is, how does the instantiation relation itself get instantiated? One might think, then, quite generally, that there is a difficulty in the idea of physical objects having intentional properties not so much because the objects are physical and the properties intentional, but simply because the objects are objects and properties properties, and because the whole idea that objects have properties -- I mean any objects and any properties -- is a confusion.

I think it should be granted that, if this is your concern, then the pleonastic conception of properties might help you; at least it should be granted that, if this is your concern, then conceivably the distinction between the pleonastic and non-pleonastic conceptions of properties is relevant to your problem. But the trouble here is that it is very difficult to see why this kind of consideration has any bearing on the problem of intentionality. That there is a problem of intentionality presupposes that there is a difficulty in the idea of a physical object's having a intentional property, but the difficulty in question can hardly be that the notion of instantiation is obscure. That difficulty does not distinguish the problem of intentionality from a perfectly general metaphysical problem about properties, namely, the problem how it is possible that, in general, objects can instantiate properties. But one can raise this
perfectly general problem without mentioning intentional properties, or even the mind, at all. One need simply ask `How it is possible that a physical object can have a physical property?’. But surely, if there is a problem of intentionality, that problem must of necessity involve the mind! It follows that if there is a reason for finding difficulty in the idea of a physical object’s having a intentional property, and if that reason gives rise to the problem of intentionality, it cannot be the reason that, in general, it is hard to see how objects instantiate properties.

And here, I think, we arrive at a general characterization of Schiffer’s mistake in supposing that very abstract issues in the metaphysics of properties have an impact on the problem of intentionality. To decide to present the problem of intentionality as a problem about intentional properties is to decide to formulate the problem within a particular metaphysical framework, the framework of properties. Whether to hold a pleonastic or a non-pleonastic conception of properties is, I think, an important question: it is a question about whether to adopt the framework of properties. And whether there is a problem of intentionality is also an important question: it is a question about whether an apparent problem, a problem formulable within the framework of properties, and quite possibly formulable within other frameworks, really is, all things considered, a genuine problem. But these two questions are quite independent of each other: a question about the framework of properties is one thing; a question about a problem formulable within that framework is quite another.20

20 Versions of this paper were read at the Australian National University and at the Massachusetts Institute of Technology; I would like to thank audiences at both universities for many helpful comments. I would also like to thank the following for comments and discussion on previous drafts: Ned Block, Andrew Botterell, Noam Chomsky, Paul Horwich, Robert Stalnaker, and Judith Jarvis Thomson.
Chapter Three
Disjunctive and Second Order Physicalism

I

In one of its simplest forms, physicalism involves the identification of mental properties with neurophysiological properties of the human brain. Thus for example, the simplest physicalism entails that being in pain is identical to being in a particular neural state; and that being itchy is identical to being in a particular (different) neural state.¹

The trouble with this simplest physicalism, however, is that it is vulnerable to the multiple realization objection. The objection draws our attention to dogs, octopi, echidnas, yaks, mollusks, and maybe even the odd robot or Martian. Surely some, and perhaps all, of these creatures can feel pain or be itchy. However, none of these creatures could have the property of being in a particular neurophysiological state — not, anyway, if we take this property as involving a specifically human brain. It follows that none of them could feel pain or be itchy. But surely octopi, mollusks and yaks can feel pain! And surely robots and Martians could feel itchy. At least there seems no reason to rule this out a priori. So it seems that philosophers who want to be physicalists must abandon the simplest version of their doctrine, and look for a different version.

Now versions of physicalism designed to accommodate the multiple realization objection have typically assumed one of two closely related forms. According to the first -- which I will call disjunctive physicalism -- mental

¹ This simplest version of physicalism is a doctrine about properties, as indeed are all the varieties of physicalism discussed in this paper. There are certainly ways of formulating physicalism which make no reference to properties, but I will set these aside.
properties are identical with disjunctive physical properties. Suppose that we
discover that being in pain is instantiated one way in humans and another way
reptiles. This fact refutes simple physicalism. But it would not refute a more
complicated version of physicalism according to which the property of being in
pain is identical not to being in a particular human neural state but rather to a
disjunctive property: being either in a particular (human) neural state or in a
particular (reptilian) neural state.

The second version of physicalism -- which I will call second order
physicalism -- appeals not to disjunctive properties, but to second order
properties, properties which involve quantification over other properties.
Second order physicalism has two parts. The first is the claim that mental
properties are identical with second order properties. Thus, for example, the
property of being in pain is on this view identical to the property of having some
property with a certain causal role. The second part is the claim that, as a matter
of fact, the only properties which have this causal role are physical properties.
This allows for the possibility of multiple realization: different properties
might be the property which has the causal role characteristic of pain or itching.
But it is also a version of physicalism: if the only properties which have the
causal roles characteristic of mental properties are physical, then "all
information is physical information". ²

Second order a... disjunctive physicalism are, as I said, closely related.
How closely related are they? A number of philosophers have recently
suggested that they are so closely related that the differences between them, if

² Second order physicalism might more correctly be viewed as a combination of functionalism
and physicalism, but I will ignore this complication here. The physicalist slogan that all
information is physical information is due to Jackson (1982). For discussion of how to make
this slogan more precise, see Lewis (1983, 1994), Jackson (1993) and Pettit (1994).
any, do not really matter. The clearest exponent of this view is Jaegwon Kim. In a recent series of papers, Kim argues that, at least in the standard presentations of these doctrines, they are simply the same doctrine: second order physicalism simply is disjunctive physicalism.

The thesis that there is no difference between disjunctive and second order physicalism -- or, more weakly, that the differences between these two doctrines do not bear emphasis -- is, if true, significant. One of the tasks of any serious philosophy of mind is to provide a statement of what the most plausible form of physicalism is. Disjunctive physicalism, however, does not seem a very plausible proposal. There are a number of reasons for this, but the one that most moves Kim is the idea that, if disjunctive physicalism is true, then mental properties will fail to be what he calls natural or genuine properties. The notion of a natural or genuine property as Kim develops it is a complex notion with many strands; however, speaking rather roughly, a natural property is a property that figures in a scientific law, a property that is causally efficacious, and perhaps a property that determines a natural kind. As Kim puts it, from the truth of disjunctive physicalism, "it follows that mental kinds are not causal kinds, and hence are disqualified as proper scientific kinds" (1993: 327).

---

3 See Kim (1993, 1994). A similar view is expressed by Hartry Field when he writes: "The differences between this nonfunctionalist physicalism and the functionalist physicalism I prefer are not really terribly important. I see functionalist physicalism as only a slight generalization of nonfunctionalist physicalism; the differences between these two doctrines have been greatly overemphasized (1992: 277). (It is clear from context that by 'functionalist physicalism' and 'nonfunctionalist physicalism', Field means what I am here calling second order physicalism and disjunctive physicalism.) David Lewis expresses a related though different view in (1983, 1994). The proper interpretation of Field and Lewis, however is a difficult matter, and I certainly do not want to simply lump their views in with Kim's. For the purposes of this paper, therefore, I will concentrate solely on Kim.

4 Terminology: As I read him, Kim tends to use the expression 'natural kind' interchangeably with expressions such as 'natural property' or 'genuine property'. However, I will speak here simply of properties and not of kinds.
Now of course the suggestion that disjunctive physicalism entails that mental properties are not natural or genuine properties would be of no consequence to physicalists if physicalists had another version of physicalism at their disposal. However, if Kim is right that disjunctive physicalism disqualifies mental properties from being natural properties, and if he is also right that second order physicalism simply is disjunctive physicalism, then presumably what goes for the latter goes also for the former. Moreover, since these are the only two forms that the thesis of physicalism might assume compatibly with multiple realization, physicalism in general will yield the result, as Kim puts it, that "psychology as a science with disciplinary unity turns out to be an impossible project" (1993:327).

The idea that there is no serious difference between second order and disjunctive physicalism is significant for a more general reason also. Any serious philosophy of mind must not only take a stand on what the best statement of physicalism is; it must also take a stand on what divides physicalism and anti-physicalism. According to those who discern no difference between second order physicalism and disjunctive physicalism, however, what divides physicalism and anti-physicalism is, among other things, the question of whether mental properties are genuine or natural properties. According to Kim, for example, the "anyone interested in defending a serious dualist position will eschew the realization talk altogether and

---

5 I assume here that there are only three varieties of physicalism: simple physicalism; disjunctive physicalism, and second order physicalism. Could it not be, however, that physicalism might take a different form: for example, could a physicalist not be token-identity theorist or a supervenience theorist? My reply is that to the extent that these latter doctrines are varieties of physicalism, they are identical to one of either simple, disjunctive or second-order physicalism, and to the extent that these doctrines are not identical to either simple, disjunctive or second-order physicalism, they are not varieties of physicalism. For discussion of these issues, see my (1994b) and Kim (1993).
consider mental properties on a par with physical properties" (1993: 314). And Kim goes on to say that the serious dualist will consider mental properties as first order properties "in their own right, characterized by their intrinsic natures" (1993: 314). The suggestion implicit in such remarks is what divides physicalists and anti-physicalists is a commitment to the status of mental properties: physicalists hold that mental properties are not genuine or natural properties; anti-physicalists hold that they are.

In this paper, however, I want to argue that Kim and the line of thought he represents is mistaken on three counts. First, Kim is mistaken not to sharply distinguish second order physicalism and disjunctive physicalism. Second, Kim is mistaken to think that physicalism is in general incompatible with the idea the mental properties are not genuine or natural. There is perhaps some reason to think that this is true of disjunctive physicalism; however, there is no reason to think it true of second order physicalism: at least by the criteria of naturalness of property that Kim offers us, second order properties certainly count as natural. Third, Kim is mistaken to construe the dispute between physicalists and anti-physicalists as a dispute between those who hold that mental properties are genuine and those who don't. This is not only because according to second order physicalism, mental properties are perfectly genuine: it is also because the best version of anti-physicalism also assumes that mental properties are second order properties.

II

According to second order physicalism, the property of being in pain is the second order property of having some property that has a certain causal role.
The first thing to be said about this view is that it certainly doesn't appear to be the same view as disjunctive physicalism. Off hand, a second order property is one thing, a disjunctive property is another.

According to Kim, however, appearances are deceptive. "On the construal of mental properties as second order properties", he says, "mental properties will in general turn out to be disjunctions of their physical realization bases" (1993:324). In particular, if

pain is the property of having a property with specification H, and if \( N_h, N_r, \) and \( N_m \) are the properties meeting H, pain turns out to be the disjunctive property \( N_h \vee N_r \vee N_m. \) If you hold the second order property view of mental properties, pain has been reduced to, and survives as, this disjunctive physical kind. (1993: 332)

In short, Kim suggests that second order physicalism provides no alternative to disjunctive physicalism because, given plausible assumptions, second order properties simply are disjunctive physical properties.

What is Kim's argument for the collapse of second order and disjunctive physicalism? The first premise is simply a statement of property identity that all second order physicalists are committed to:7

\[
(1) \quad \text{The property of being in pain} = \text{the property of having some property that has causal role C.}
\]

The second premise of the argument is a further statement of property identity, this time identifying the second order property with a disjunctive property. Suppose that the property that actually has causal role C is the disjunctive

---

6 Strictly speaking, one should say that instantiations or occurrences of properties have causal roles not the properties themselves, but I will ignore this complication.

7 Following standard practice, I will in what follows call the human neural property \( N_h \) and the reptilian neural property \( N_r, \) and call the causal role definitive of being in pain C.
property of being either in a human neural state or in a reptilian neural state. Then, says Kim, the second order physicalist is committed to the claim that:

(2) The property of having some property that has causal role $C =$ the property of being either in $N_h$ or in $N_r$.

From these two premises it clearly follows that:

(3) The property of being in pain = the property of being either in $N_h$ or in $N_r$.

But of course (3) is simply a statement of disjunctive physicalism. In short, then, Kim argues that the second order physicalist is committed to both (1) and (2), and that the conjunction of these entail disjunctive physicalism.

But if this is Kim's argument for the collapse of second order physicalism into disjunctive physicalism, it is open to two important objections. The first is that it is far from clear why we should accept the truth of (2). It is clear of course that the two properties mentioned in (2) are coextensive in the actual world. That is, if we assume (as we are assuming) that the only things which actually have pain are humans and reptiles, then the property of having some property that has causal role $C$ and the property of being either in $N_h$ or $N_r$ are coextensive. But this observation cuts no ice. The actual coextension of these properties no more establishes their identity than does the actual coextension of being a cordate and being a renate establish theirs.

It might be suggested that (2) is true because the properties are not only co-extensive in the actual world, but are co-extensive across all possible worlds. However -- while co-extension across possible worlds is certainly a plausible account of property identity -- it is very doubtful that the two properties are co-extensive across all possible worlds. After all, it is not hard to imagine possible worlds in which objects have the property of having some property
that has causal role $C$, and yet do not have the disjunctive property referred to in the right hand side of (2). But if there are such worlds, then the two properties are not co-extensive across all possible worlds, and they are not identical.

More plausibly, it might be suggested that if there is a nonactual possible world in which something is in pain with out having $N_h$ or $N_r$ but instead has (for example) property $A$, then, in place of (2), at friend of Kim's argument should simply advance (2'):

\[(2')\quad \text{The property of having some property with causal role } C = \text{ the property of having either } N_h \text{ or in } N_r \text{ or in } A.\]

Of course, (2') would not yield (3) when conjoined with (1); but it would yeild a parallel conclusion, namely:

\[(3')\quad \text{The property of being in pain = the property of being either in } N_h \text{ or in } N_r \text{ or in } A.\]

Moreover, it might be said, (3') is as good as (3) for Kim's purposes, since (3') is a statement of disjunctive physicalism just as (3) is.

However, there is a serious difficulty in pursuing the strategy of substituting (2') and (3') for (2) and (3). Let us suppose that there is a property $A$ which in some nonactual possible world $W$ plays the causal role definitive of being in pain, namely $C$. The trouble with this suggestion is that it seems quite possible that this very property might be instantiated in the actual world \textit{without} playing $C$; in other words, it is possible that someone in the actual world might have this very property without being in pain. But this means that (3') is false; for if (3') is true, it is impossible for someone to have $A$ and yet not be in pain. To illustrate, suppose that $A$ is the property of being wet. We can

\[8\text{ I am indebted here to a discussion with Judith Jarvis Thomson.}\]
imagine a possible world in which being wet plays the causal role definitive of pain. Still, it is quite obvious that there are things in the actual world which are wet and yet not in pain. It follows that being in pain is not identical to the disjunction being wet or in Nh or Nr. For if it were, anything wet in the actual world would be in pain, which is patently false. More generally, it follows that (3') is false, and moreover, since the argument from (1) and (2') to (3') is valid, that (2') is false too. In other words, it is of no help to Kim to replace (2) and (3) with (2') and (3').

What is it about (2) that makes it attractive to Kim? Kim suggests that the truth of (2) is made plausible by a certain view about the semantics of second order property designators, singular terms of the form "The property of having some property that has causal role C". He says:

we may in general consider second order property designators as nonrigid. (1994: 9)

That is, we may in general consider a second order property designator as denoting different properties in different possible worlds. This view certainly does give Kim a reason to believe the truth of (2). If expressions such as "the property of having some property that has causal role C" are nonrigid designators, we could take them as referring to different properties in different possible worlds. Kim suggests that, in this world, "the property of having some property that has causal role C" nonrigidly denotes the disjunctive property, being either in Nh or Nr. In another possible world it may denote a different property. If this view of the semantics of second order property designators were correct, then, we might derive (2), and from it and (1), derive (3).
However there is every reason to be suspicious of the doctrine that second order property designators are nonrigid.\(^9\) Consider *first order* property designators, such as 'the property of being round'. Expressions of this type are plausibly rigid; 'the property of being round' is plausibly a rigid designator of a property, the property of being round. The reason for this might be described in the following way. When we consider that the extension of the predicate 'is round' is different in different possible worlds, we are not imagining that objects in these different possible worlds have different properties. We are rather imagining that different objects in different possible worlds have the same property. In short, it seems that 'the property of being round' is a rigid designator.

The trouble for Kim, however, is that what goes for first order property designators goes, *mutatis mutandis*, for second order property designators. When we consider that the extension of the predicate 'has some property that has causal role C' is different in different possible worlds, we not imagining that different objects have different properties. *A fortiori*, we are not imagining that the objects in these possible worlds have different physical properties. Rather we are imagining that different objects have the same property. And this suggests again that second order property designators are rigid. However, if these expressions are rigid, then there is no reason to accept (2) as true.

There is also another and more direct way to scotch the idea that second order property designators are non-rigid. For consider the paradigmatic cases of nonrigid designators, viz., descriptions of the form 'the President of the United States', 'my cat', 'A girl from the north country', and so on. What is

\(^9\) For further discussion of rigid designation and property designators, see Jackson (1980a), and Prior et al (1982). I am very much indebted to both papers, and to discussion with Judith Jarvis Thomson, for the next three paragraphs.
typical of all these expressions is that they are semantically equivalent to expressions which involve quantification. For example, at least according to the standard Russellian view, `the President of the United States' is semantically equivalent to `there is a unique $x$ such that $x$ is the President of the United States'. On the other hand, however, second order property designators (and indeed, this is true of property designators quite generally) do not seem to be semantically equivalent to expressions which involve quantification. What, for example, is the quantifier expression which is semantically equivalent to `the property of having some property that has causal role $C$'? That there is no candidate expression is very good evidence that second order property designators are rigid designators.

To summarize, then, our first objection to Kim's argument for the identification of second order and disjunctive physicalism is that there is no reason to believe, and much reason to doubt, the truth of his premise (2). Perhaps it will be replied to our first objection, however, that, whether or not there is good reason to accept the truth of (2), the physicalist must accept that (2) is true, for otherwise there would be nothing to his physicalism. After all, there is nothing physicalist about (1) taken alone. Taken alone, (1) merely says that the property of being in pain is a second order property with a particular specification. But that is perfectly consistent with the falsehood of physicalism. In short, a physicalist must hold something in addition to (1) to qualify as a physicalist, and (2) seems the best option.

But this reply merely raises our second objection to Kim's argument. The second objection to Kim's argument is that, while it is perfectly true that the second order physicalist is committed to something more than (1), it is by no means obvious that he or she is committed to (2). For it is important to
distinguish two things. On the one hand, there is the property of having some property that has causal role C. On the other hand, there are the properties that have causal role C. Now (2) says that the property of having some property that has causal role C is identical with a disjunctive physical property. In fact, however, the second order physicalist is not required to identify the first of these properties with a physical property at all, still less with a disjunctive property. What requires defence is rather the idea that the properties which have causal role C are as a matter of fact physical properties. In the case that we are considering, what this amounts to is the claim the properties that have causal role C are being in Nh and being in Nr. At most, then, the second order physicalist needs to defend:

(4) The properties that have causal role C are the property of being in Nh and the property of being in Nr.

But (4) is quite different from (2). In particular, the conjunction of (1) and (4) does not entail (3) as the conjunction of (1) and (2) does. In other words -- and this is our second objection to Kim's argument -- while it is perfectly true that the second order physicalist is committed to (1) and also that (1) and (2) entail disjunctive physicalism, the second order physicalist is not committed to (2). In place of (2), the second order physicalist can suggest (4), and thereby avoid the collapse of his position into disjunctive physicalism.

III

Thus far, then, we have distinguished second order physicalism from disjunctive physicalism, and have taken a dim view of Kim's argument that these are at bottom the same doctrine. First, the actual coextension of disjunctive and second order properties is no indication that they are identical.
Second, disjunctive properties and second order properties are not coextensive across possible worlds because the expressions referring to these are plausibly taken are rigid. Finally, the second order physicalist is not in any case obliged to identify disjunctive properties and second order properties.

I want now to turn from the question of the distinction between disjunctive and second order physicalism to Kim’s more general view that if physicalism is true, then mental properties cannot be natural or genuine properties. Kim offers two different arguments for this view: the projectibility argument, and the explanatory exclusion argument.

At the beginning of the paper, I remarked that the notion of a natural property is a complex notion with many strands. The projectibility argument selects one strand in the natural property notion: the idea that a property is natural if it figures in a scientific law. In particular, the projectibility argument purports to show that physicalism is incompatible with the existence of psychological laws, and, moreover, since mental properties are the properties which figure in such laws, physicalism is incompatible with the naturalness of mental properties. As we will see, however, there is a weakness in the projectibility argument, and this is that it overlooks the distinction between second order and disjunctive physicalism.

The projectibility argument proceeds as follows. Let us imagine that the property of being in pain is, as a matter of fact, realized in, and only in, humans and reptiles. It follows from disjunctive physicalism that being in pain is identical with having either a human brain in a certain state or having a reptilian brain in a certain state. Now let us suppose further that the statement, (5) Everything that is in pain winces.
is a psychological law. If we suppose that (5) is a law, then it follows, given
disjunctive physicalism, that a further statement, namely,
(6) Everything that is either in N₁ or N₂ wincing
is also a law.

On the other hand, (6) is not a law. A necessary condition on something’s being a law is that it is projectible, where by `projectible' Kim means that the law is confirmed by observations of its positive instances (1993: 319-20). For example, if `All men are mortal' is a law, it follows that it is subject to confirmation by the observation of positive instances, namely, by the observation of mortal men. If, somehow or other, `All men are mortal' were not confirmed by the observation of mortal men, we should conclude that it is not projectible and, therefore, is not a law.

But the difficulty with (6) (the projectability argument proceeds) is that it is not projectible; it is easy to imagine cases in which it is not subject to confirmation by positive instances. Imagine -- for example -- that we go back and examine all the records of past observations on the basis of which we have supposed (6) to be a well-confirmed law. To our surprise, we discover that the only cases we have considered were of humans wincing, and no reptiles wincing. Of course, all these examples of wincing humans were positive instances of (6). But we cannot suppose that they are confirming instances of (6). Certainly these wincing humans confirm another law nonidentical to (6), namely `Everything that has a human brain in a certain state winces'; but this latter law is not (6). So (6) is not confirmed by its positive instances. However, if it is not confirmed by its positive instances it is not projectible. And if it is not projectible, then not a law. However, if (6) is not a law, then neither is (5).
Conclusion: (5) is not a law, and moreover it is hard to imagine how, if disjunctive physicalism were true, there could be any psychological laws at all.

Now, in his discussion of this argument, Kim assumes that it applies not only to disjunctive physicalism, but to physicalism generally. In addition, he argues that the conclusion of the argument is not so damning to physicalism as might be initially supposed. In particular, Kim says, the physicalist who lives without psychological laws, and who thereby admits that mental properties are not natural, is not thereby committed to eliminativism, which he understands as the doctrine that there are no pains, no beliefs and so on just as there are no "phlogiston, witches and magnetic effluvia" (1993: 334). According to the physicalist, Kim says, the ontological status of pains, beliefs and the rest is rather like the ontological status of tables: "Tables do not constitute a scientific kind. There are no laws about tables as such, and being a table is not a causal-explanatory kind. But that must be sharply distinguished from the false claim that there are no tables. The same goes for pains" (1993: 334).

Kim is of course quite correct in saying that the conclusion of the projectibility argument does not entail eliminativism, at least as he understands that doctrine; however, that does not mean that this conclusion is not unacceptable on other grounds. For consider: if the projectibility argument is correct, psychological theories must have a neural or behavioural vocabulary and cannot have what might be called a mentalistic vocabulary, where a mentalistic vocabulary is either our normal intentional and qualitative vocabulary, or else is a technical descendent thereof. The reason for this is simply that if the projectibility argument is right, then any statement formulable in mentalistic vocabulary will fail to be a law. Surely, however, it would be a very bad thing if physicalism had the consequence that psychological
theories must have a certain vocabulary. After all, the question of what is the appropriate vocabulary for psychological theories is an empirical issue. Of course, it might turn out that the appropriate vocabulary for psychology is a neural vocabulary, but there is no particular reason to suppose that it will. Physicalists who argue that psychology must have a certain vocabulary appear to be adopting the unfortunate position of illegitimately letting metaphysics dictate to empirical science.

Given the distinction between second order and disjunctive physicalism however, there is no need for the physicalist to take up the position of letting metaphysics dictate to science. It is crucial to the success of the projectibility argument that if (5) is a law, (6) is a law. And I think we should agree that if disjunctive physicalism is true, then it is the case that if (5) is a law (6) is. On the other hand, however, if second order physicalism is true, there is no reason to agree that if (5) is a law (6) is. In particular, if second order physicalism is true, from the fact that (6) is not a law it does not follow that (5) is not. It is therefore perfectly possible for the second order physicalist to agree that (6) is not a law. For admitting this is not admitting that (5) is not a law. In short, then, the projectibility argument does not show that second order physicalism is incompatible with the existence of psychological laws; as a consequence, the projectibility argument does not show that physicalism in general is inconsistent with the naturalness of mental properties.¹⁰

¹⁰ Ned Block (1995) argues that projectibility argument might be developed on weaker premises than those involved in Kim's discussion. According to Block, it is not crucial to the argument that (5) is logically equivalent to (6). What is crucial is rather that (5) and (6) are nomologically equivalent, i.e., true in the same class of nomologically possible worlds. The controversial aspect of Block's position is whether nomological equivalence is enough to guarantee that if (5) is a law, (6) is, but this is an issue that I cannot discuss here. Therefore my criticism of the projectibility argument is subject to the important qualification: that it is a criticism of the argument only to the extent that the latter is interpreted as presupposing that (5) and (6) are logically equivalent.
In suggesting that the projectability argument overlooks the distinction between second order and disjunctive physicalism, I do not mean to suggest that the argument might not also be in trouble on other grounds. On the contrary, given the unclarity of the notions in which the argument is framed -- the notions of projectability, lawhood, disjunctive property and so on -- it seems to me unclear whether the argument establishes that (6) is not a law and that disjunctive properties are not natural.\textsuperscript{11} Nevertheless, it does seem worthwhile to point out that the argument fails at least because of the difference between disjunctive and second order properties. The reason for this is that, quite apart from the projectability argument, there appears to be an impressive consensus that disjunctive properties are unnatural because they are unfit for laws. My point in this section is only that there is nothing in second order physicalism which precludes the second order physicalist from agreeing with this consensus.

VI

To this point I have argued, first, that there is a clear distinction between second order and disjunctive physicalism and, second, that the distinction is important because it allows physicalists to defeat the projectibility argument against disjunctive physicalism.

I turn now to the second argument that Kim advances in favour of the view that mental properties are not natural or genuine: the explanatory exclusion argument. The explanatory exclusion argument takes up another strand in the notion of a natural property: the idea that if a property is natural, then it is causally efficacious. According to this argument, if physicalism is

\textsuperscript{11} For some argument along these lines see Block 1995.
true, then mental properties are not natural, because they are not causally efficacious.

The explanatory exclusion argument begins by considering the following three claims:

1. John winces because he is in pain.
2. John winces because he is in a particular human neural state.
3. The fact that John is in pain is not identical to the fact that John is in a particular human neural state.

It seems very plausible to suppose that the second order physicalist is committed -- or might be committed -- to the truth of each of these claims. Commitment to (9), for example, is a central commitment of any version of physicalism; and commitment to (8) is just common sense. In addition, second order physicalism is committed to (10) because, according to this position, the property of being in pain is not identical to the property of being in a particular human neural state. In short, then, the second order physicalist is committed to (8)-(10). However, says Kim, the problem for the second order physicalist is that it is impossible to be committed to the joint truth of (8)-(10).

Why is it impossible to be committed to the joint truth of (8)-(10)? Kim argues that the truth of (8)-(10) is inconsistent with a principle which he calls the principle of explanatory exclusion. Kim says:

The general principle of explanatory exclusion states that two or more complete and independent explanations of the same event or phenomenon cannot co-exist. (1993: 250)

---

12 I assume here that the causal relata are facts rather than Davidsonian events. As far as I can tell, nothing turns on this assumption and equivalent points would arise in the Davidsonian framework. For further discussion of the explanatory exclusion argument, see the papers by Davidson, Kim and McLaughlin in Mele and Heil 1993.
According to Kim (8)-(10) provide an instance in which one event or phenomenon has two complete and independent explanations. For (8) and (9) jointly assert that two facts are explanations for one fact, the fact that John winced. And (10) asserts that the fact that John is in pain is not identical to the fact that John is in a particular neural state. This suggests that these two facts are complete and independent explanations of one fact. But that is inconsistent with the principle of explanatory exclusion.

How should one react to the apparent inconsistency of the principle of explanatory exclusion and (8)-(10)? Kim argues that the only response available to the second order physicalist is to deny the truth of (8). (9) cannot be given up without compromising physicalism; and (10) cannot be given up without abandoning second order physicalism. Thus, if the principle of explanatory exclusion is to be maintained, the only possibility is to reject (8). Moreover, if (8) is false it is hard to see how any equivalent claim about the causal efficacy of mental facts or properties could be true. Kim concludes that, if second order physicalism is true, then (8) and any equivalent claim is false; he summarizes this by saying that mental properties are not causally efficacious.

How successful is the explanatory exclusion argument? The first thing to be said about this argument is that it doesn't suffer from the weakness of the projectibility argument. As we saw, the weakness in the projectibility argument was that it overlooked the distinction between disjunctive physicalism and second-order physicalism. The explanatory exclusion argument, on the other hand, would, if successful, apply equally to both varieties of physicalism. For it is plain that disjunctive physicalism is also committed to (8)-(10): (8) and (9) are true for the disjunctive physicalist for the same reason that they are true for the second order physicalist; and (10) is true,
because the property of being in $N_h$ is not identical to the disjunctive property of being either in $N_h$ or in $N_r$. As far as the explanatory exclusion argument is concerned, then, what goes for second order physicalism goes also for disjunctive physicalism.\(^{13}\)

Even if the explanatory exclusion argument applies to both varieties of physicalism, however, there is nevertheless a crucial weakness in the argument, and this is with the principle of explanatory exclusion itself. There is certainly some initial plausibility to the idea of explanatory exclusion; however, on reflection, I think the principle ought to be rejected. The problem is as follows. If the principle of explanatory exclusion applies as Kim's argument requires it to apply, then it will entail that no paradigm case of causal efficacy is a case of causal efficacy. However, a principle about causal efficacy which entails that no paradigm case of causal efficacy is an extremely implausible principle.

To illustrate this, let us first take note of two points. First, the explanatory exclusion principle, and the argument founded on it, does not merely apply to the particular case represented by (8)-(10). For there is nothing special at all about (8)-(10); on the contrary, if the explanatory principle is correct, it will yield the result that any three statements of similar structure cannot be true together. Moreover, if Kim is right that (8) is to be rejected, then the explanatory exclusion argument will yield the result that no multiply realized property at all could play a role in causation, or could be causally efficacious.

\(^{13}\) Since the explanatory exclusion argument applies to disjunctive physicalism and to second order physicalism, I will for brevity focus my discussion of the argument solely on second order physicalism.
Second, it is very implausible to suppose that multiple realization is only a feature of the relation between psychological and neurological properties. Most likely, neurological properties are multiply realized by chemical properties; and chemical properties are multiply realized by physical properties. Perhaps also physical properties themselves will be multiply realized. And so on. Indeed, it seems at least an epistemic possibility that any paradigm case of a causal efficacious property will turn out to be multiple realized. In short, for any property at all, if it does turn out to be multiply realized, then, by Kim’s principle, it can be established that the property in question is not causal efficacious.

But now we can see why the explanatory exclusion principle is implausible. For quite generally it is very implausible to suppose that no paradigm case of causal efficacy or of causation is a case of causal efficacy or of causation. Suppose, for example, someone were to offer an account of the colour blue according to which nothing which we take to be blue was blue (not the sky, not the sea, etc.) and the only thing that was blue was some object on Alpha Centauri which nobody had seen. If it is agreed -- as I think it should be -- that such an account is an implausible account of the colour blue, then it should equally be agreed that the explanatory exclusion principle is implausible. A principle about causation or of causal efficacy must at least yield that some of our paradigm cases of causal efficacy are cases of causal efficacy. It is a decisive consideration against any account of causation or causal efficacy that incorporates the principle of explanatory exclusion that it fails to do this. 

14 I owe this example to Ian Gold.
In short, then, the explanatory exclusion principle has an unacceptable consequence, namely, that it might be that no paradigm case of a causally efficacious property is causally efficacious. This suggests -- to return to our main line of argument -- that Kim is quite mistaken in thinking that the only response to the apparent inconsistency of (8)-(10) and the principle of explanatory exclusion is to reject (8). A completely different, and perhaps more reasonable, response is to keep (8) but modify the principle so as it is not inconsistent with (8)-(10). And indeed, when we examine the principle of explanatory exclusion closely, it becomes obvious that such a modification is eminently possible. For in fact it is not very clear what the principle of explanatory exclusion says. As Kim himself notes, the difficulty in understanding the principle is in understanding precisely what 'complete' and 'independent' are to mean. Kim himself offers no general account of what these are supposed to mean. What he does instead is describe a range of cases intended to be illustrative of the idea of two or more complete and independent explanations, one of which is the case illustrated by (8)-(10). 16 However, if the only way we have of understanding the principle of explanatory exclusion is to examine cases to which it allegedly applies, then, I suggest, we ought to interpret the principle as not applying to cases of the structure of (8-10). For as

---

15 Though he doesn't put matters in quite this way, I take the argument here to follow Yablo (1992). For further discussion of the importance of paradigm cases in conceptual analysis, see Johnston (1992).

16 Kim says: "The meanings of 'complete' and 'independent' are obviously crucial. I shall not be offering definitions of these terms, rather I shall focus on some specific cases falling under the intended distinctions with the hope that, in the course of my discussion, reasonably determinate core meanings will emerge that will give the exclusion principle clear and substantial enough content" (1993: 250). It is interesting to note however that in the paper from which this quotation is drawn -- 'Mechanism, Purpose and Explanatory Exclusion', reprinted as chapter 13 of his 1993 -- Kim did not regard cases of the structure of (8-10) as falling under the principle. The extension to these cases comes later. See, in particular, Kim 1993, chapters 17 and 18, as well as Kim 1994.
we have seen, if the principle is interpreted so as to apply to those cases, it is an extremely implausible principle. On the other hand, of course, if the explanatory exclusion principle is interpreted so as not to apply to (8)-(10), then explanatory exclusion argument collapses.

VII
So far I have been concerned mainly with three arguments of Kim's about second order physicalism. The first argument was that second order physicalism collapsed into disjunctive physicalism. I replied that this argument illegitimately assumed that second order property designators are non-rigid. The second argument was the projectibility argument, which purported to show that second order physicalism entailed that there are no psychological laws because any putative such laws would not be projectible. I replied that while disjunctive physicalism might have this result, second order physicalism did not. The third argument was the explanatory exclusion argument, which purported to show that second order properties could not be causally efficacious. I replied that the principle behind this argument -- the explanatory exclusion principle -- was in Kim's presentation of it too strong to be plausible, and, moreover, that if the principle is weakened, it ceases to have the result that second order properties are not causally efficacious.

There is one last matter that emerges from Kim's discussion of second order properties that I want to consider. Kim suggests that

the serious dualist will, or should, take mental properties as genuine first order properties with their own distinctive intrinsic natures, whereas the functionalist...considers mental properties as..."second
order' constituted by their causal roles and specified by job descriptions (1993: 364).

In this and similar passages, Kim seems to be articulating a certain conception of what divides the physicalist from the anti-physicalist in the philosophy of mind. According to this conception, what divides physicalist from anti-physicalist is "the propriety and significance" (Kim 1993: 316) of mental properties: according to the physicalist, mental properties will be either disjunctive or second order properties, and this means that they will be nongenuine or nonnatural properties; according the anti-physicalist, on the other hand, mental properties will be neither disjunctive nor second order, and this means that they will be genuine or natural properties.

However, I think that there are two reasons for rejecting this account of the debate between physicalists and anti-physicalists. First, as we have seen, the best version of physicalism -- namely, second order physicalism -- does not have the result that mental properties are not genuine or natural. As we saw in section II, the best version of physicalism is committed to claims (1) and (4), here repeated as (11) and (12):

(11) The property of being in pain = the property of having some property that has causal role C.

(12) The properties that have causal role C are the property of being in Nh and the property of being in Nr.

This view does entail that mental properties are second order properties. But it does not entail the idea that mental properties are disjunctive properties, or that they cannot figure in scientific laws, or that they are not causally efficacious. Hence there is no reason to think that it entails that mental properties are not
genuine or natural -- not anyway, if we are working with the standards of
naturalness that Kim offers us.

But there is a second reason to reject Kim's account of the debate
between physicalists and anti-physicalists. Kim seems to suggest that that
apparatus of second order properties is either unavailable to antiphysicalists, or
at least should be rejected by them. However, it is hard to see why this should
be so. According to one particularly obvious version of anti-physicalism, for
example, (11) is true, but (12) is false, for the reason that the properties which
has causal role C is not some physical properties but rather some properties of
an irreducibly mental substance, perhaps a soul. According to this kind of
view, mental properties will certainly be second order properties, because
statements of property identity such as (11) will be true. Indeed, it is somewhat
curious that this kind of position is apparently neglected by Kim, since one of
the original motivation for claims such as (11) is precisely that (11) is consistent
not only with physicalism, but also with anti-physicalism.

Perhaps it will be replied to this, however, that the potential existence of
the theory I have just imagined is no objection to Kim, because Kim is only
interested in the kind of view held by serious antiphysicalists. Talk of souls, it
might be said, is not serious. Fair enough; but in any case there seems to me to
be a different and much more serious version of anti-physicalism which talks
not of souls but of brains, and entails that mental properties are second order
properties.

The kind of view I have in mind is constructed to ape second order
physicalism as much as possible. In particular, in place of (11) and (12), the
anti-physicalist position I am imagining would suggest (11*) and (12*):

(11*) The property of being in pain = The property of having
some property that has causal role C and is pain presenting.

(12*) There is some property P such that P is a property of the brain,
P has causal role C and and is pain presenting.

Here, (11*) is intended as a somewhat stronger claim than (11). Anti-
physicalists often object that claims such as (11) must be false, because, if (11)
were true, one could be in pain and not have the peculiar phenomenal quality
associated with pain.\textsuperscript{17} We might put this by saying that, if (11) is true, one
could be in pain, and yet the pain would not be \textit{pain-presenting}, that is, it would
not present itself as painful to the subject of the pain. (11*) is intended to
remedy that defect in (11): according to (11*) the property of being in pain is
identical with a property which not only has a certain causal role, but also
which is pain-presenting.\textsuperscript{18} Similarly, (12*) is intended as a somewhat weaker
claim than (12). Perhaps it is true that in order to be taken seriously, anti-
physicalists must agree that the mind is identical to the brain. If so, however,
the anti-physicalist is perfectly free to accept (12*) in place of (12): (12*) says
nothing about \textit{what} property of the brain is at issue, so it is consistent with (12*)
that the property which has causal role C and is pain presenting is a non-
physical property of the brain.\textsuperscript{19}

In sum, then, if the sophisticated physicalist is understood as holding
(11) and (12), and if the sophisticated anti-physicalist is understood as holding

\textsuperscript{17} For a classic statement of this objection, the so called absent qualia objection against
physicalism, see Block (1980).

\textsuperscript{18} The idea of pain-presentingness is suggested by Jackson (1980a), though he uses it in a
different way.

\textsuperscript{19} I assume here that there is some way of saying what counts as physical property (or as a
non-physical property) of the brain. This is obviously a tententious assumption, but one that I
think can be defended in a limited kind of way. For some discussion of this issue, see Crane
and Mellor 1990 and my 1994a.
(11*) and (12*), then it is plain that Kim’s construal of the debate about physicalism is mistaken. Both positions are committed to mental properties being second order properties, and both positions are committed to mental properties being realized in the brain.

If the debate about physicalism is not properly understood as a debate about the status of mental properties, how is it to be understood? The best way to see what is at issue between physicalist and anti-physicalist, I think, is to see what considerations influence the choice between (11) and (12) on the one hand, and, (11*) and (12*) on the other. And here the discussion exhibits a subtle but familiar oscillation between matters empirical and matters conceptual. On the one hand, for example, the anti-physicalist might argue that conceptual considerations push us toward (11*) and (12*) rather than (11) and (12). After all, from a certain conception of what a mental property is, it seems obvious that (11) leaves out something out, whereas (11*) does not, and thus that the latter should be preferred over the former.

This conceptual consideration is weighty, but not, I think, decisive. For it is open to the physicalist to respond that empirical considerations push us to adopt (11) and (12). The physicalist might perhaps argue that what we currently know about the brain rules out the truth of (12*). Or, more generally, the physicalist might argue that we have good empirical reasons for believing physicalism, and, since (11*) and (12*) are not compatible with physicalism, we therefore should believe (11) and (12). As for the conceptual considerations suggesting (11*) over (11), the physicalist might reply that, while (11*) certainly appears true from a naive point of view, physicalism has taught us something
about the proper conception of the property of being in pain; in particular, it has taught us that (11) is true, and not (11*).20

My own view is that the physicalist is here overstating things, and that a sober assessment of the current state of our knowledge of the brain, combined with a reflection on our conception of what a mental property is, suggests very strongly that the rational choice is (11*) and (12*) over (11) and (12). But my purpose here has not been to argue for this.21 My purpose rather has been to demonstrate that Kim and the line of thought he represents misconstrues not only the differences between the varieties of physicalism but also the debate between physicalists and anti-physicalists.22

20 This position, or something like it, seems to be that of David Lewis. See his 1995.

21 For some argument to this effect, see my 1995a.

22 For comments on previous drafts and for discussion of these topics, I should like to thank: Ned Block, Andrew Botterell, Jennifer Noonan, Joe Lau, Rob Stainton, Jason Stanley, Robert Stalnaker, Zolani Szabo, and Judith Jarvis Thomson.
There are at least two ways to criticise the doctrine of physicalism in the philosophy of mind. The first way is to argue that physicalism is false. An example of such an argument is the argument from qualia. On the one hand (proponents of the argument from qualia argue), there is something it is like for me smell a rose and this something-it-is-like is a qualitative aspect of my experience of smelling a rose. On the other hand, a purely physical object in a purely physical world is such that, for it, there could not be something it is like to smell a rose. Hence physicalism is false; for, according to physicalism, I am a purely physical object in a purely physical world.

The other way to criticise physicalism is to question, not the truth of the doctrine, but its content. According to physicalism, I am a purely physical object in a purely physical world. However, what can it possibly mean to say that I am a 'purely physical' object in a 'purely physical' world? These phrases might seem clear; certainly, philosophers of mind go around using them as though they were clear. But in fact they are not clear. Moreover, without a definite idea of what 'physical' means, we can have no idea of what physicalism is. According to this second way of criticizing physicalism, then,

---

1 The doctrine of physicalism that I will discuss sometimes goes by other names -- 'materialism', for example, or 'naturalism' -- however I will here confine myself to 'physicalism'. This is in part simply for brevity, but it is also to emphasize the limits of my discussion. Both 'materialism' and 'naturalism' have connotations beyond the philosophy of mind: 'materialism' is historically associated with a position in theology, viz. atheism; and 'naturalism' is historically associated with a position in moral philosophy, viz., the reduction of moral facts or properties. In a full scale examination of the doctrine of physicalism one would of course have to discuss these matters, but my focus will be solely on philosophy of mind.
physicalism is not true, not so much because it is false, but because it is too unclear to have a truth value at all.

It is a curious fact that, apart from some notable exceptions\(^2\), this second kind of criticism has not been seriously discussed much in philosophy of mind. And this is not because the unclarity in the word `physical' and similar expressions has gone unnoticed. On the contrary, this unclarity is in many cases explicitly noted, only to be set aside. Frank Jackson, for example, writes:

> A fair question is how to specify precisely the notion of a physical property. I am not going to answer this fair question. Roughly, I will mean what is typically meant: the kinds of properties that figure in, or are explicitly definable in terms of, those that figure in physics, chemistry, biology and neuroscience. This rough characterisation leaves it open why those sciences rather than say psychology or politics, are chosen to settle the favoured class, and it says nothing about how committed this approach is to those sciences being roughly right in the kinds of properties they need for their own internal purposes. Nevertheless, I think that the rough characterization will do for our purposes here. As far as I can see, nothing in what follows turns on the answers to these controversial matters. (1994: 26)

The suggestion implicit in such remarks is not, I think, that `physical', `material' and their cognates are perfectly clear. The suggestion rather is, first, that mere unclarity does not by itself constitute an objection to physicalism, and, second, that there is no compelling reason to think that the unclarity in its central notions can be developed into an objection against physicalism. For one thing,
Jackson seems to be suggesting, one can if pressed fairly easily give an account of what `physical' means -- of what a physical property is\(^3\) -- and so that any apparent unclarity in these notions can be eliminated. For another, even if there were some unclarity in the notion of a physical property, nothing would turn on it: the unclarity of physicalism does not much matter to the main agenda of philosophy of mind.\(^4\)

I will argue, however, that the unclarity of physicalism does matter to philosophy of mind. To begin with, I would like to distinguish three different ways that one might develop this unclarity into an objection against physicalism, as follows:

(1) **The Empty or False Objection**: The word `physical' and its cognates suffer from an ambiguity. In one way of resolving the ambiguity, physicalism is an empty doctrine; in the other way of resolving the ambiguity, physicalism is obviously false. Hence physicalism is either empty or obviously false.

(2) **The Triviality Objection**: The word `physical' and and its cognates are meaningful; but their meanings do not license a distinction between something that is physical and something that is not physical. Hence physicalism is trivial.

---

\(^3\) Physicalism is sometimes associated with the repudiation of abstract objects such as properties, possibilia, chances and the like. However, I intend no such restriction and will follow Jackson in supposing that physicalists face the fair question of saying what a physical property is. For some discussion of the connection between physicalism and the repudiation of abstract objects, see Crane and Mellor 1990 and Stoljar 1995a.

\(^4\) The idea that the unclarities of physicalism can be ignored is a common theme in Jackson. In `Epiphenomenal Qualia', for example, he says "I take the question of definition [of physicalism] to cut across the central problems I want to discuss" (1982: 469). And in an earlier paper, he remarks "I will take for granted the notion of a physical property. It is explained variously in various places; but nothing in what follows hangs on the differences between the explanations, and I take the general idea to be familiar enough" (1980: 26). It is obvious also that Jackson is here speaking for many.
(3) The No Reason Objection: The word 'physical' and its cognates have various meanings; but under no reasonable interpretation is there as much reason to believe physicalism as is commonly supposed. Hence while physicalism might be nontrivially true, there is little or no reason to believe that it is true.

To the extent that the unclarities of their doctrine have been discussed at all by physicalists, the overwhelming tendency, I think, has been to concentrate on the first two of these objections. My main concern here, however, is not so much to consider these first two objections, or to criticize the answers that physicalists have provided to them. My aim rather is to suggest that even if those answers to the first two objections are correct, they leave the physicalist with no convincing answer to the third.

II

In a recent series of papers, Tim Crane and D.H. Mellor argue that physicalism, which they understand as the doctrine that everything is physical,\(^5\)

faces a dilemma: 'physical' is either taken in a very restrictive sense, in which case physicalism is clearly false; or it is taken in a very broad sense, in which case the doctrine is almost empty. The challenge to the physicalist is to define a doctrine that is both defensible and substantial. (Crane 1993: 224)\(^6\)

---

\(^5\) According to Crane and Mellor, physicalists "believe that everything is physical...that all entities and properties, relations, and facts are those which are studied by physics and other physical sciences" (1990: 82). There are a number of problems with this definition, but these problems will not be our main concern here. For a more adequate way to spell out physicalism, see Lewis (1983, 1994), Kim (1993), and Pettit (1994).

\(^6\) Crane is here summarizing the argument presented more fully in Crane and Mellor (1992). See also Hempel 1970.
This is what I have called the Empty or False Objection. According to Crane and Mellor, one way of defining physicalism results in a doctrine that is obviously false; another way results in a doctrine that is empty.

What are the two senses of 'physical' that Crane and Mellor have in mind? In the first sense, 'physical' applies to a property or fact which is best described or explained by some ideal or future physics, perhaps the physics that the best human scientists can devise some time far in the future. However, the problem with defining 'physical' this way is that physicalism now seems quite empty; that is, it now seems a doctrine that anyone, including philosophers traditionally regarded as anti-physicalists, might believe. After all, if present day physics proves incapable of explaining some phenomena -- the phenomena of qualia, for example -- then there is nothing to rule out the possibility that it will simply be extended in the future to do so. But this means that qualia become physical by default, and there is no longer any question about the physicalistic status of qualia.

In the other sense intended by Crane and Mellor, 'physical' applies to a property or fact which is best described or explained by contemporary physics. However, the problem with defining 'physical' this way is that physicalism

7 The restriction to physics is unnecessary here. One might just as easily have spoken of 'the physical sciences', where that is taken to mean physics, chemistry, biology and neuroscience. Compare the quotation from Jackson with which I began.

8 One might distinguish here between ideal physics and future physics. After all, it is not a trivial thesis that the physical sciences which our descendants will produce far in the future will explain everything about our world. The reason for this is that we are plausibly regarded as epistemically bound: there are limits on the things that we as a species can understand. So the physics of even our most sophisticated descendants might very well fail to explain everything. An ideal physics, on the other hand, presumably would explain everything -- that, indeed, is what would mark such a physics as ideal. The issues of epistemic boundedness -- the phrase is Jerry Fodor's -- are interesting, and interact in interesting ways with the unclarities of physicalism we are discussing, but for the most part I will not discuss them. For general discussion of epistemic boundedness, see Chomsky (1975), and Fodor (1983); for discussion of the impact of epistemic boundedness on philosophy of mind, see the final pages of Jackson (1982), as well as Fodor (1983: 120-126), and McGinn (1991).
now seems obviously false. Crane and Mellor provide two reasons for this. First, if 'physical' is to be defined in terms of present physics, then physicalists are committed to the idea that "any future extensions of [present day physics] would not be physics: that physics...is already complete. But nobody believes this" (1992: 85). Second, if physicalism is to be defined in terms of contemporary physics, then physicalists are committed to the idea that mental properties, for example, are reducible to contemporary physical properties. However, they argue, "reducibility in practice is neither feasible nor to the point" (1992: 88). In short, then, the Empty or False Objection poses the following dilemma for physicalists: if 'physical' is defined with reference to a future or ideal physics, then physicalism is empty; but if 'physical' is defined with reference to contemporary physics, then physicalism is obviously false.

How successful is the Empty or False Objection? At first sight it seems persuasive. An unadulterated appeal to future or ideal physics certainly would result in physicalism's being an empty doctrine. And nobody thinks that physics as it is right now is complete. However, there is a way in which a physicalist might avoid the dilemma, and this is appealing to contemporary physics on the one hand, but agreeing that contemporary physics is not complete on the other.9

Let us note to begin with that this is in fact what contemporary physicalists do: "I take it that physicalism is metaphysics built to endorse the truth and completeness of physics more or less as we know it" (Lewis 1986: x). Lewis, then, is at least one contemporary physicalist who defines his doctrine with the help of contemporary physics. However, does Lewis believe what

---

9 I emphasize here that in arguing that physicalist might answer the Empty or False Objection, I am not arguing that physicalists can answer all objections that have their source in the unclarity of physicalism. On the contrary, I will argue that the best answer to the Empty or False Objection leaves the physicalist with no answer to the No Reason Objection.
Crane and Mellor say that \textit{no one} believes, namely, that contemporary physics is complete? Not quite. Lewis does not claim that physics as it actually is right now is completely true. He claims rather that it is not unreasonable to regard a theory that bears some close resemblance to contemporary physics as being completely true. And this crucial emendation shows how physicalists might define their doctrine with reference to contemporary physics in the face of Crane and Mellor's remark that no-one believes that physics as it is right now is complete. A physicalist can agree with Crane and Mellor that no-one believes that physics as it is right now is complete. On the other hand, a physicalist can argue that it is by no means \textit{obviously false} that a theory which bears a strong resemblance to contemporary or near contemporary physics is completely true. And this is enough to answer the Empty or False Objection.

A somewhat different and complementary way for a physicalist to defend the completeness of present physics was suggested in a paper by the physicist Gerald Feinberg some years ago (1966), and has been developed by J.J.C. Smart (1978; 1986). According to Feinberg, while there may well be revolutionary changes in physics, these changes are likely to occur not in the physics of bulk matter, but rather in the physics of elementary particles. Feinberg goes on to say that "the two fields are almost disjoint", and moreover, that there is a good sense in which the contemporary physics of bulk matter is complete: that is, as far as bulk matter is concerned, contemporary physics has solved what Feinberg calls \textit{The Thales Problem}, the problem of saying of what kind of stuff the world is made (1966: 11).

The interest in Feinberg's remarks for the physicalist is obvious. Of course one cannot deny that there might be revolutionary changes in physics, and consequently, that there might be changes in what kind of properties are to count as 'physical'. However, one can argue that, for the philosophy of mind,
all that seems to be relevant is the physics of bulk matter, and hence that likely changes in future physics are irrelevant for the philosophy of mind. As Smart puts it, for example, the present day physicalist can tie the "definition of physicalism to present day physics, and the twenty first century physicalist can tie it to twenty first century physics. There will be no important difference which is relevant to the physicalist theory of mind, because neurons are 'ordinary matter', and the physics of ordinary matter is essentially complete and is unlikely to change in important respects" (1978: 340).10

Now it might be replied to both of these suggestions that near contemporary physics, or the physics of bulk matter, however entrenched they are at the moment, might in the future nonetheless turn out to be false, and thus that the physicalist who appeals to either is at least being extraordinarily optimistic about the present state of our knowledge.11 However, this observation, while perfectly true in itself, is not such as to affect the physicalist's reply to the Empty or False Objection. In order to answer the Empty or False Objection, it is not required that physicalism be true; still less is it required that physicalists give up their optimism about present science. All that is required is that physicalism is not obviously false. And it seems clear

10 In his defense of the idea that physicalists might appeal to contemporary physics, Lewis also cites Feinberg. See Lewis 1994. Indeed, one might argue that Lewis must appeal to Feinberg for the reason that Lewis does not specify the respect in which contemporary physics must resemble near contemporary physics for it to be the case that physicalism is nontrivial. The Feinberg-Smart view, on the other hand, does give some account of this resemblance: for physicalism to be nontrivial, contemporary and near contemporary physics must be identical with respect to the physics of bulk matter, although they might differ with respect to the physics of elementary particles.

11 The extreme epistemic optimism of physicalism is emphasized by Jackson. He writes: "It is not sufficiently appreciated that physicalism is an extremely optimistic view of our powers. If it is true, we have, in very broad outline admittedly, a grasp of our place in the scheme of things. Certain matters of sheer complexity defeat us...but in principle we have it all" (1982: 476).
enough that one could agree with Lewis, Feinberg and Smart without holding a view that is obviously false.

So much, then, for the first reason that Mellor and Crane provide for not appealing to contemporary physics -- the reason that no one thinks that contemporary physics is complete. What of the other reason that they provide -- that reduction to contemporary physics is out of the question? The problem with this dismissal of reduction is that it underestimates the sophistication of the reductive strategies that physicalists have at their disposal. There are two rather different reductive strategies that physicalists may employ, a simple strategy and a sophisticated strategy. The simple strategy is to argue that mental properties are identical with neurological properties, or possibly with properties discussed by other physical sciences. Thus, for example, according to this kind of view, the property of being in pain is identical to the property of being in such and such a neural state. The sophisticated strategy is to argue, first, that mental properties are identical to second order properties which admit of topic-neutral specification and, second, that these second order properties are realized by first order physical properties. Thus, for example, according to this kind of view, the property of being in pain is identical to the (second order) property of having some property with such and such a causal role. Moreover, according to this kind of view, the property which has this causal role is, as a matter of contingent fact, the property of being in such and such a neural state.

Now Crane and Mellor's dismissal of reduction perhaps has some force if one has only the simple reductive strategy in mind. Considerations of multiple realization are plausibly thought to rule out the idea that properties such as being in pain might simply be identified with neurological properties.\(^\text{12}\)
However, things are rather different if one has the sophisticated reductive strategy in mind. Here, considerations of multiple realization bear no weight at all; the sophisticated strategy easily allows for the multiple realization of mental properties and states. As against Crane and Mellor then, it doesn't seem out of the question that physicalists might carry through their reductive program so long as it is granted that that program might employ the sophisticated method of reduction.

It might be objected that there are reasons for doubting that even the sophisticated reductive strategy can be completed. Once again, however, while it may very well be true that there are reasons for thinking that the sophisticated reductive strategy cannot be completed, this is quite irrelevant to the Empty or False Objection. In order to answer this objection, the physicalist does not need to answer every objection to physicalism: all that needs to be shown is that it is not obvious that a physicalism defined in terms of contemporary or near contemporary physics is false. However it is perfectly

12 But for a dissenting view, see Kim 1993 Ch. 16.

13 Close examination reveals that many recent naturalistic theories of intentionality conform to what I am here calling the sophisticated reductionist model. Jerry Fodor, for example, remarks that, according to the kind of theory of intentionality he favours, "all that matters for meaning is 'functional' relations (relations of nomic covariance) between symbols and their denotations. In particular, it doesn't matter how the covariation is mediated: it doesn't matter what mechanisms (neurological, intentional, spiritual, psychological, or whatever) sustain the covariation"(1992: 56; Emphasis in original).

14 One might reply that the sophisticated method of reduction is not a method of reduction, and hence that a physicalist who appeals to this method is not a reductive physicalist in Crane and Mellor's sense. I disagree with this, but in any case it is irrelevant: whether or not the sophisticated method of reduction really is a method of reduction, the physicalist can nonetheless appeal it to answer Crane and Mellor's objection that reductionism is implausible. For discussion of the connection between reductionism and physicalism, and for criticism of the idea that physicalism can be held in a nonreductive form, see Kim (1993) and Stoljar (1995a; 1995b)
clear that, even if the sophisticated reductive strategy cannot be completed, it is not *obvious* that it cannot be completed. 15

III

At the heart of the Empty or False Objection is the idea that the physicalist is uncomfortably caught between an allegiance to contemporary physics and an allegiance to future or ideal physics. The next objection that I want to discuss -- the Triviality Objection -- develops the idea that physicalism is unclear in a slightly different way. According to this objection, the problem for physicalism springs, not from an allegiance to contemporary or future physics, but from an allegiance to the false physics of the past.

A statement of the Triviality Objection is most easily found in some of the writings of Noam Chomsky. Chomsky begins with the observation that the mind-body problem certainly made sense at the time of Descartes:

The mind-body problem can be posed sensibly only insofar as we have a definite conception of body. If we have no such definite and fixed conception we cannot ask whether some phenomena fall beyond its range. The Cartesians offered a fairly definite conception of body

15 I have assumed in this section that physicalists will respond to the Empty or False Objection by appealing to contemporary physics. However, it is worth emphasis that not all contemporary physicalists appeal to contemporary physics to define their doctrine. Philip Pettit, for example, defends the idea that physicalism is the doctrine that "actually (but not necessarily) everything non-microphysical is composed out of microphysical entities" (1994: 1). Pettit says that a physicalist of this sort "may be more or less sanguine about the accuracy of actual physics, or even about the propriety of its methods: he may be more or less optimistic about how far actual physics is on the right track" (1993: 214). Another attempt at defining physicalism which makes no reference to contemporary physics appeals to the idea that physicalism is true just in case the sciences adequate for explaining inorganic phenomena are adequate for explaining everything (cf. Block 1980; 296). These proposals are interesting and raise difficulties of their own, but for our purposes they can be set safely aside. The existence of ways to define physicalism which make no appeal to contemporary physics does not, I think, affect the main point I want to make, which is that physicalists can answer the Empty or False and Triviality Objections, but not the No Reason Objection.
in terms of their contact mechanics, which in many ways reflects common understanding. Therefore they could sensibly formulate the mind-body problem [...]. (1988: 142)

However, Chomsky goes on to say, the conception of body associated with 'contact mechanics' was rejected by the scientists following Descartes:

The Cartesian conception of a second substance was generally abandoned in a later years, but it is important to recognize that it was not the theory of mind that was refuted (one might argue that it was hardly clear enough to be confirmed or refuted). Rather, the Cartesian concept of body was refuted by the seventeenth-century physics, particularly in the work of Isaac Newton, which laid the foundations for modern science. Newton demonstrated that the motions of the heavenly bodies could not be explained by the principles of Descartes's contact mechanics, so that the Cartesian concept of body must be abandoned. (1988: 143)

If the Cartesian concept of body was abandoned, what conception of body took its place, in particular, what conception of body, or of the physical, is contemporary science operating with? The correct view, Chomsky suggests, is that "we simply abandon the whole conception of body as something possibly distinct from something else and use the methods of rational inquiry to learn as much as we can about the world -- what we call the material world, whatever exotic properties it turns out to have" (1988: 146). He concludes that "the mind-body problem remains the subject of much controversy, debate, and speculation, and in this respect the subject is very much alive. But the discussion seems to me incoherent in fundamental respects" (1988:146).

The crucial idea in these remarks, I think, is the idea that it is mistaken to suppose that we have a conception of body as something possibly distinct from
something else. At least since Newton, Chomsky is suggesting, we have had no such a conception. Now this does not mean, of course, that words like 'body', 'material' and 'physical' do not currently have meanings at all. It is rather that their meanings do not license a distinction between the physical world and the non-physical world. More exactly, to say that a thing is part of the physical world is not to prejudice its nature; it is simply to say that it is part of the world, i.e., it is simply to say that the thing exists.  

Now from this point of view, the thesis of physicalism is entirely trivial, and for a very simple reason: its denial is incoherent. In particular, it is incoherent to say, as Frank Jackson for example says, that "certain properties of certain mental states, namely those I've called qualia, are such that their possession or absence makes no difference to the physical world" (1982: 473). Jackson seems here to be distinguishing the physical world from the non-physical world rather as if these were two kinds of boxes, and then supposing that, while qualia exist, they belong in the non-physical box. According to the notion of the physical world at issue in the Triviality Objection, however, there is no question of saying both that qualia exist and that they are not part of the physical world. For if qualia exist, they are ipso facto part of the physical world. According to this conception of the physical therefore, even Jackson is a physicalist. Hence physicalism is trivial.

16 In discussion, Chomsky has suggested that the expression 'the physical world' should be understood via an analogy to the expression 'the real truth'. It would obviously be an absurd misunderstanding to suppose that the expression 'the real truth' marks one kind of truth, real truth from another kind of truth, non-real truth. Similarly, according to Chomsky, it is a misunderstanding to suppose that there are two kinds of world, a physical world and a non-physical world.

17 It is important to emphasize here that the incoherence that the Triviality Objection finds in Jackson's position has nothing whatsoever to do with qualia themselves. The incoherence rather has its source the property that Jackson says qualia have, viz., existing and yet not being part of the physical world.
How successful is the Triviality Objection? I think physicalists should immediately concede that, in a certain sense, Chomsky is right. Physicalists should immediately concede, that is, that there is a sense of the phrase `the physical world' in which, if qualia exist, then they are part of the physical world. On the other hand, however, it is not obvious that physicalists cannot reply to the Triviality Objection by simply digging in their heals and insisting that, while there may be various understandings of the expression `the physical world,' the understanding that is at issue in physicalism is rather different from the one Chomsky describes.

After all, it does not seem on the face of it impossible for physicalists to avail themselves of conception of the physical which would allow us to meaningfully assert that qualia (say) are exist but are not part of the physical world. Now of course this would not mean reverting to the Cartesian conception of body. What it might mean, however, is introducing such a notion by stipulation. There is nothing easier than delineating a certain class of properties and dubbing these the physical properties. That would be enough to answer the Triviality Objection: stipulating what class of properties are physical would result in a perfectly meaningful definition of physicalism. Or, to take a somewhat less artificial strategy, one might answer the Triviality Objection in the same way that one might answer the Empty or False Objection, viz., by appealing to contemporary, or near contemporary, science. We saw in our discussion of the Empty or False objection that contemporary physicalists can appeal to contemporary physics -- and the other physical sciences -- to provide a family of properties reasonably called physical. Once such a family of properties is set out, one might perfectly well ask whether particular mental properties lie beyond its range.
It might be replied that these responses miss an important rhetorical effect of the Triviality Objection. The figure of Descartes casts a very long shadow over contemporary philosophy of mind. Indeed, much philosophy of mind can be seen, as Gilbert Ryle saw it in the 1940s, as an attempt to explode Descartes's myth of the ghost in the machine. It is therefore important to be reminded, as Chomsky's Triviality Objection does remind us, that in fact there is nothing `spooky' or `antiscientific' about Descartes's problem of mind and body and the science that generated it: on the contrary, his scientific methodology is perfectly respectable, it is simply predicated on a false theory.18

But it is no part of my purpose to deny the rhetorical effect behind the Triviality Objection. Indeed, I want to emphasize this rhetorical effect, for it is precisely this effect that will be exploited by the No Reason Objection, the objection which I think is the strongest way to develop the unclarities in physicalism. But none of this affects the fact that the physicalist can answer the Triviality Objection. To answer the Triviality Objection, all that is required is the possibility of physicalism's being nontrivially true; it is not at all required that either physicalism or its denial be, in Chomsky's phrase, "offensive to 'scientific common sense'" (1988: 48).

IV

Thus far we have considered two ways to develop the idea the physicalism is unclear: the Empty or False Objection, and the Triviality Objection. What unites these two objections is that they concentrate on the statement of physicalism: according to both there is no reasonable way to state physicalism

18 Compare Philip Pettit, who remarks that, according to physicalism, "the actual microphysical facts fix the actual non-microphysical facts under the contingency, perhaps inter alia, that there is no spooky stuff around" (Pettit 1994; 3; Emphasis added).
according to which it might be defended as nontrivially true. I think it is time to allow that these objections fail, and that physicalism might be stated in such a way that it might be true. However, granting this much is not granting that there is no problem for physicalism as far as the unclarity in `physical' is concerned. According to the final objection I want to discuss -- the No Reason Objection -- the central problem raised by the unclarity of these notions does not concern the statement of physicalism: the central problem concerns our reason to believe physicalism. In particular, according to the No Reason Objection, one would only think that one had much reason to believe physicalism if one had failed to notice the ambiguities in `physical' that we have been discussing.

At least two major reasons are often cited in support of the belief that physicalism is true. The first is the idea that to renounce physicalism is, as Stephen Schiffer has put it, to "renounce the scruples of the natural scientist" (1987:142). For Schiffer, then, one believes physicalism because commitment to physicalism is part and parcel of a commitment to the scientific world view. If physicalism goes, then so too does the scientific world view.

I think it is quite obvious that Schiffer is speaking here for many physicalists; much of the attraction of physicalism does come from the vague idea that `science shows physicalism to be true'. However, I think it is just as obvious that one would only find Schiffer's reason for adopting physicalism persuasive if one had failed to notice the ambiguities in physicalism that we have been discussing. In one sense, `physical' applies to a property or fact which some ideal future physics is in the best position to describe. In another sense, `physical' applies to a property or fact which contemporary physics is in the best position to describe. From our earlier discussion of the Empty or False Objection, it is clear that, as far as philosophy of mind is concerned, what is at
issue is this second, narrower, sense of 'physical'. However, if physicalism is to be interpreted this way, it is quite clear that to give it up is not to give up the scruples of the natural scientist. In this way of understanding it, to give up physicalism is merely to accept that contemporary physics -- or, more broadly, the class of sciences which include contemporary physics, biology, chemistry, and neuroscience -- is not up to the task of explaining *everything* about the world, or about our minds. But this is hardly to give up the scruples of the natural scientist. Indeed, just the reverse is true: one might argue that the scruples of the natural scientist are most prominently displayed, not in the remark that contemporary science explains everything about the world, but rather in the remark that the current state of our knowledge is in many respects limited and fragmentary.

The second reason for believing physicalism derives not from the idea that to give up physicalism is to give up the scientific world view, but rather from the idea that certain very plausible principles *within* science would be jeopardized if physicalism were not true. Hartry Field, for example, argues that one should believe physicalism because without it we are left without an explanation of what he calls the "mesh between psychology and the lower-level sciences", that is, the fact that psychological explanations and (say) neurophysiological explanations, while potentially in conflict with one another, apparently do not conflict (1992: 285). Field writes:

> Without a quasi-reductionist account of the mesh in explanations between sciences, it looks as if this mesh is a total mystery. To a large extent, it is because we should avoid believing in total mysteries that we ought to assume that if a special scientific theory is accepted, then a quasi-reduction to lower level sciences is possible. (1992:286).
According to Field, then, to give up physicalism is to give up the only available account of the mesh between the sciences. We need a way of understanding this mesh, and physicalism provides the only way to do this.

Now Field is certainly right that an account of the mesh between various sciences is required; and he is also right that the reductionist model he sketches provides an account of this mesh. Nevertheless, I think it is mistaken to assume that these points provide much reason for believing physicalism. For one thing, there might well be another account of this mesh; indeed, Field himself readily admits that this is a possibility (1992: 286). Moreover -- and this reason is more germane to our discussion -- attention to the unclarities of physicalism reveals a significant difference between physicalism on the one hand, and the reductionist model that Field is defending on the other.

According to the reductionist model that Field has in mind, "for each sentence in the language of a successful special science, like chemistry or genetics or psychology, there is a sentence in the language of a lower level science -- and ultimately in the language of physics -- that in some intuitive sense `expresses the same facts’" (1992: 272). There are a number of things to be said about this model of reduction, but for our purposes it is sufficient to note the following. A commitment to this kind of reductionism is not yet a commitment to physicalism, for the simple reason that Field’s model of reduction does not specify what the language of physics is. It is clear from our discussion of the Empty or False Objection, of course, that physicalists must assume that what is at issue in the reductionist model is the language of contemporary or near contemporary physics. But the present point is that there is nothing in the model of reductionism itself which dictates this. Field’s reductionist model entails that for every sentence in the language of psychology, there is a sentence in the language of physics which expresses the
same facts. But physicalism entails something stronger: it entails that for every sentence in the language of psychology, there is a sentence in the language of contemporary, or near contemporary physics, which express the same facts.

Once it is appreciated that commitment to Field's model of reductionism does not entail commitment to physicalism, the possibility opens up that one may hold the model of reductionism without holding physicalism. And indeed, given our previous discussion, it is not hard to imagine what such a position would be like. One might hold, for example, that the model of reductionism is perfectly correct but that the language of physics needs to be expanded or modified to ensure that for each sentence in (say) the language of psychology, there is some sentence in the language of physics which, in Field's phrase, in some intuitive sense 'expresses the same facts'. Or, to take a somewhat different possibility, one might argue that if there really is a potential conflict between two sciences, one might always avoid that conflict by revising the lower level science. Field's model of reduction is perfectly compatible with the idea that, to explain the mesh between various sciences, it is necessary, not to modify the upper science, but to modify the lower-level science.

The point being made here is simply a generalization of a point often made in discussions of reductionism. In discussions of reductionism it is often pointed out that the basic notions of higher level sciences might have to be changed or modified in order to ensure that these sciences can be reduced to lower level sciences. What is less often pointed out is that this point generalizes to the basic notions of lower level sciences: that is, the basic notions of lower level sciences might have to be changed or modified in order to ensure that higher level sciences can be reduced to them. But this opens up a possibility that is not apparent in Field's discussion: that one might accept his model of reductionism, and yet not be a physicalist, because one might be confident that
new concepts or explanatory principles will be required to achieve the reduction.\textsuperscript{19}

In short, then, what goes for Schiffer’s reason for believing physicalism goes also for Field’s. The reasons provided by both philosophers lose much of their persuasiveness if one pays attention to the unclarities of physicalism. In the case of Schiffer, to get clear about physicalism is to get clear that to giving up physicalism is not giving up the scientific world view. In the case of Field, to get clear about physicalism is to get clear that commitment to the reductionist model of the mesh between the sciences is not yet commitment to physicalism.\textsuperscript{20}

\textsuperscript{19} It is important here to emphasize that, at least on some views of the history of science, the revision of lower-level sciences is more common than the revision of upper-level sciences. Chomsky remarks, for example, that "most of the history of science is like that, at least as I read it. To the extent that unification has been achieved, it has not been achieved through strict reductionism, except in rare cases. It’s very commonly been the case that what we think of as the more fundamental science had to be radically revised" (1993: 86). For further discussion, see Chomsky 1995.

\textsuperscript{20} There is a third major reason -- deriving from considerations of mental causation -- for thinking that materialism is important which is in some respects similar to Field’s. Tyler Burge summarizes these considerations as follows: "Appeal to mental causation that does not consist in physical causation appears...to require us to doubt the adequacy of current forms of physical explanation, even within the physical domain. So such an appeal ought to be rejected" (1992: 114-15). Burge himself does not find these considerations persuasive. He writes "The reasoning seems to me to have some force. But I think that it is not as forceful as it may appear. Why should mental causes alter or interfere with the physical system if they do not materially consist in physical processes? Thinking that they must, surely depends on thinking of mental causes on a physical model -- as providing an extra ‘bump’ on the effect....But whether the physical model of mental causation is appropriate is...part of what is at issue (1992: 115). I am in sympathy with Burge here, but to adequately defend the position he outlines would take us deeply not only into the metaphysics of causation, but also into the question of whether mental causation does indeed conflict with the laws of classical physics such as the conservation of energy law. Both of these topics go beyond the scope of what I can attempt in this paper; hence my conclusions in this section must be tentative. For the classical presentation of the causal argument for physicalism, see Davidson 1970. For presentations of the argument which improve in Davidson’s in various ways, see Schiffer 1987 chapter 6 and Kim 1993 chapters 17 and 18. For criticism of the Davidson-Schiffer-Kim position, see Burge 1992. For criticism of the idea that mental causation would violate the conservation of energy law, see Averill and Keating 1981.
It might be replied, however, that our discussion of the reasons for believing physicalism is misguided. After all, the No Reason Objection is quite different from the Empty or False and the Triviality Objections. Unlike these latter objections, the No Reason Objection is not attacking the idea that physicalism could be nontrivially true. However, if the No Reason Objection is not attacking the possible truth of physicalism, why can a physicalist not defend the truth of physicalism even while granting the No Reason Objection, and admitting that to give up physicalism is not to give up any general philosophical principles that cannot reasonably be given up? More briefly, a physicalist might simply reply that what matters for physicalism is that it is true, not that we have overwhelming reason to believe it.

There is certainly some truth to this reply; nothing in the No Reason Objection precludes the idea that physicalism could be true, or that one might meaningfully debate its truth. Nevertheless, I think we should insist that philosophers who believe physicalism must provide some justification for why they do so beyond the mere assertion that they do so.

The central reason for insisting on this is that many problems in philosophy of mind have a structure which presupposes that there is a lot of reason to believe physicalism. In the typical case, for example, one is first presented with the thesis of physicalism, and then invited to agree that physicalism is inconsistent with our naive conception of the mind. A solution to the problem involves seeing to what extent one might retreat from the naive conception of mind in order to 'save' the mind for physicalism.21

21There are many examples of philosophers who conceive of things this way. Here is Hartry Field, for example, describing the problem he calls Brentano’s problem: "[this is] the problem (which Brentano believed to be unsolvable) of giving a materialistically adequate account of the believing, desiring, and so forth. Unlike Brentano, I take it is unquestionable (given what we know about the world) that materialism is true; I also take it as unquestioned for purposes of this paper that people do believe and desire. These two assumptions together amount to the
Now of course different philosophers react to problems of this structure in different ways. Some think that no retreat is possible, and therefore infer from the fact that our naive conception of mind is inconsistent with physicalism that no-one has any minds; others infer that our naive conception must be mistaken in various ways, and therefore propose physically respectable conceptions to take its place. For our purposes, however, the crucial consideration is not to what extent one can retreat from our naive conception of mind. The crucial consideration is rather the following: unless there is persuasive reason to believe physicalism, we will have no reason to retreat from our naive conception at all. If physicalism is not a persuasive doctrine, that is to say, then the fact, if it is a fact, that the our naive conception of mind is inconsistent with physicalism will provide no reason at all for revision of that conception.

To put the point another way, the physicalist is asking us to concentrate on a particular class of properties, the physical properties, and then argues that it is crucial to the significance and status of mental properties that they be members of this class. However, the question to put to the physicalist is the following: what is so special about that class of properties? There are obviously infinitely many classes of properties, some of which mental properties are members of, some of which mental properties are not members of, and some of which it is not clear either way. But why should it matter to the assumption that Brentano's problem can be solved; what I shall be interested in is the question of how it can be solved, that is, of what the assumption that it can be solved shows about the nature of belief and desire" (1978: 70-1 Emphasis in original).

A third way to react of course is by deciding that physicalism is false. However, amongst those who take physicalism seriously, this is the opinion of a small and rather sad minority. Frank Jackson, for example, says that the assumption that physicalism is true is "an assumption I take, reluctantly, to be contrary to fact" (1994: 26; Emphasis added). It is worth mention also that even while Jackson gives up physicalism, he clearly sees the interest in discussions of physicalism as deriving from the question of to what extent the existence of minds is consistent with physicalism; in short, antiphysicalists such as Jackson also assume that physicalism is important.
significance and status of mental properties, or to the having of mental properties, that they be members of one class rather than another? Of course, the physicalist might well offer reasons for thinking that the class of physical properties is favoured in some way. For example, it might be argued with Schiffer that the scruples of the natural scientist dictate that one advert only to the physical properties in our theories of the world. Or, again, it might be argued with Field that one loses our only reasonable account of the mesh between various sciences if one mentions properties not on the physicalist's list. As we have seen, however, it is precisely these reasons that the No Reason Objection calls into question.23

Finally, it is worth emphasis that what is true of the No Reason Objection would also be true of the Empty or False and Triviality Objections if those objections are sound. I have been arguing that, if the No Reason Objection is correct, then physicalism gives us no reason to revise the naive conception of mind. But of course the same thing is true if physicalism is either trivial or false: after all, if physicalism is false, then clearly it is of no consequence that the naive conception of the mind is inconsistent with it; and if physicalism is trivial or empty, then there is no reason to think that it will be inconsistent with the naive conception. On the other hand, however, it is important to see that the No Reason Objection has a significant advantage over either the Empty or False or Triviality Objections, and this is that it can be sound, even if these other objections are unsound. Indeed, not only is it true that the No Reason Objection can be sound even if these other objections are unsound; it is also true -- at least it is true if our previous argument is correct.

23 The caveat mentioned in footnote 20 is worth reiteration here: I have not discussed one major reason for thinking that materialism is important, namely the reason having to do with mental causation.
that the best answers that physicalists can provide to the first two objections I distinguished, leave them with no convincing answer to the third. In short, even if we grant physicalists everything, that will not be enough.

In summary, then, the central moral of our discussion of the unclarities of physicalism is not that physicalism is false, or that it cannot help being true, or that it lacks a truth value; the central moral is that there are not, as is commonly supposed, many persuasive reasons to believe its truth. And this means that the air of intellectual anxiety that pervades many discussions in philosophy of mind is misplaced. It may be that in the future course of science, mental properties are assimilated to the class of physical properties in the way that Lewis, Smart Field and many others suppose; or it may be that they resist such assimilation in the way that contemporary critics of physicalism such as Jackson supposes. Either way, however, it is not clear that philosophers of mind should care very much about the outcome.24

V

I began by distinguishing three ways in which one might develop that idea that physicalism is unclear. The first two ways resulted in the Empty or False and the Triviality Objections, and I argued that these objections were answerable by the physicalist. The third way resulted in the No Reason Objection, and I argued that this objection is not so answerable. In particular, the main reasons

24 I take myself to be in agreement with Burge here when he says: "It seems to me that philosophers should be more relaxed about whether or not some form of materialism is true. I think it a thoroughly open -- and not very momentous -- question whether there is any point insisting that mental events are, in any clear sense, physical" (1992: 117). For a somewhat similar view, see Stich 1992. However, both Burge and Stich arrive at their views on the basis of considerations independent of the ones discussed in this paper.
for investing physicalism with philosophical importance are only compelling if one has failed to notice the ambiguities in `physical' and similar expressions.

I want now to close the paper by considering a line of thought that runs contrary to the drift of the discussion so far. To this point the discussion resembles nothing so much as a recipe for unemployment so far as philosophers of mind is concerned. If they are not to take the problems of physicalism seriously, then what is that philosophers of mind do? What becomes of philosophy of mind?

Now the first thing to be said about this is that even if it were right, it would not be an objection. It is surely not the task of philosophy of mind to justify philosophy of mind. Moreover, philosophical arguments always carry with them the threat, or promise, that the problem under discussion will turn out to be chimerical: that is just the way of philosophical arguments. But in any case, it does not seem to me that the objection (if that is what it is) is right. That is, even if philosophers of mind ought not to engage in discussion of physicalism, it doesn't follow that they ought not to engage in discussion at all.

One way to bring this out it by recurring to the Jackson remark I quoted at the beginning. Jackson said, near enough: nothing turns on the answer to the question of what a physical property is. In the body of the paper, I have interpreted this remark as suggesting that the unclarities in physicalism are not such as to affect the doctrine. Under this interpretation, I think the remark is false: the unclarities do matter to physicalism, for ironing out those unclarities establishes that physicalism is not as important as Schiffer and Field (and, I think, many others) suppose.

But there is a somewhat more radical way to interpret the remark. Under this interpretation, the suggestion is, not that nothing turns on the unclarities in physicalism, but rather that nothing turns on physicalism itself.
According to this second and more radical way of interpreting Jackson's remark, the real issue is not, as I have been assuming, that physicalism might be inconsistent with the naive conception of mind. The real issue is rather that the naive conception of mind is, or seems to be, inconsistent with itself.

The line of thought which prompts this radical interpretation is most forcefully presented, not by Jackson himself, but by Thomas Nagel. Nagel writes:

We are not faced only with the problem of the relation between mind and body, or the inclusion of the mental in the physical world. The broader issue of personal and impersonal, between subjective and objective, arises also for a dualist theory of mind. The question of how one can include in the objective world a mental substance having subjective properties is as acute as the question how a physical substance can have subjective properties. (1979: 201)

Nagel goes on to say that all of the central disputes concerning physicalism, all "the disputes over causal role, theoretical identification, functional realization, while of interest in themselves, fail to give expression to what makes the mind-body problem so hard" (1979: 202).

There are obviously many ideas raised by this and similar remarks of Nagel's. What is important for our purposes, however, is the distinction between two kind of problem. The first is the problem of physicalism: this is the problem, as I stated it earlier, of seeing how far one needs to retreat from the naive conception of the mind in order to have a conception of mind that is consistent with physicalism. The second is the problem of objectivity: this is the problem of seeing how far one needs to retreat from the naive conception of the mind in order to have a conception of mind which is consistent *simpliciter*. I have argued that we should not take the first kind of problem seriously,
because physicalism on its own gives us no reason to retreat from the naive conception of mind. However, I have not argued -- and nor does it follow from what I have argued -- that there is no reason to take seriously the second kind of problem. What needs to be appreciated here is simply that the philosophy of mind presents us with various kinds problems each of which needs to be debated and discussed on its own terms.

25 I do not mean to suggest here that the objectivity issue is the only issue that one might discuss without discussing physicalism. I can think of two additional issues (there may be others). One concerns the description and analysis of the notions used in actual psychological theory -- notions such as computation, representation, innateness etc. -- and also the question of how these psychological notions will mesh with the notions of other sciences, and with notions drawn from common sense, in particular folk psychological notions. The other concerns the description of folk psychology itself. Recent psychological literature suggest very strongly that humans beings acquire, and acquire early, folk psychology in the normal course of cognitive development (see Fodor 1992, Wellman and Gelman 1992). However, it remains very much an open question what the principles of this acquired psychology are; in particular, it remains an open question whether the acquired psychology contains principles of instrumental rationality, principles of charity, principles of humanity as various philosophers have speculated that folk psychology contains. I concentrate on the objectivity issue in the text in part because it seems to me quite obvious that the two issues just mentioned can be discussed without discussing physicalism, and in part because the question of objectivity seems more like a legitimate heir to the question of physicalism, and so worthy of special attention.

26 I am very much indebted to Noam Chomsky for discussion on these issues and for comments on previous drafts. Thanks also to: Ned Block, Andrew Botterell, Alex Byrne, Ned Hall, Jennifer Noonan, Simon Saunders, Jason Stanley, Robert Stalnaker, Zoltan Szabo, Judith Jarvis Thomson and Steven White.
Bibliography:


Kim, J. (1994): 'Mental Causation and Two Conceptions of Mental Properties'. Ms.


*Philosophical Perspectives*, Vol.4, Ridgeview.


Reprinted in Block 1980.


Clarendon


