

This paper appeared in *Philosophical And Phenomenological Research* Vol. LXII, No.2, March 2001, 253-281. Reprinted in Chalmers, D. (ed.) 2002: *Philosophy of Mind: Contemporary and Classical Readings*, Oxford University Press. Reprinted in Nagasawa, Y, Ludlow, P, and Stoljar, D. (ed.) 2004. *There's Something about Mary: Essays on Frank Jackson's Knowledge Argument*, MIT Press.

Two Conceptions of the Physical

Daniel Stoljar

Abstract: The debate over physicalism in philosophy of mind can be seen as concerning an inconsistent tetrad of theses: (1) if physicalism is true, a priori physicalism is true; (2) a priori physicalism is false; (3) if physicalism is false, epiphenomenalism is true; (4) epiphenomenalism is false. This paper argues that one may resolve the debate by distinguishing two conceptions of the physical: on the *theory-based conception*, it is plausible that (2) is true and (3) is false; on the *object-based conception*, it is plausible that (3) is true and (2) is false. The paper also defends and explores the version of physicalism that results from this strategy.

1. One way to view the contemporary debate in philosophy of mind over physicalism is to see it as being organized around an inconsistent tetrad of theses. These are:

- (1) If physicalism is true, a priori physicalism is true.
- (2) A priori physicalism is false.
- (3) If physicalism is false, epiphenomenalism is true.
- (4) Epiphenomenalism is false.

It is obvious of course that these theses *are* inconsistent: (1) and (2) entail that physicalism is false, while (3) and (4) entail that it is true. Barring ambiguity, therefore, one thing we know is that one of the theses is false.

On the other hand, each of the theses has powerful considerations, or at least what seem initially to be powerful considerations, in its favor. [\(1\)](#) In support of (1) are considerations of supervenience, articulated most clearly in recent times by Frank Jackson and David Chalmers. A priori physicalism is a thesis with two parts. The first part—the physicalist part—is that the mental supervenes with metaphysical necessity on the physical. The second part—the a priori part—is that mental truths are a priori entailed by physical truths. Many philosophers hold that supervenience stands in need of justification or explanation; Jackson and Chalmers argue that the project of justifying or explaining supervenience *just is* the project of making it plausible that there is an a priori entailment of the mental by the physical. This suggests that the first part of

a priori physicalism inevitably involves the second. By considerations of supervenience, therefore, (1) is true: if physicalism is true, a priori physicalism is true.[\(2\)](#)

In support of (2) are considerations of the apparent epistemic distinctness of qualia from anything physical. According to many philosophers, knowledge of every physical property a person has cannot by itself suffice to know which qualia, if any, his or her experiences instantiate. [\(3\)](#) The conclusion drawn from this is that a priori physicalism is false; for if physical truths a priori entail mental truths, one could know qualia merely on the basis of physical knowledge. By considerations of epistemic distinctness, therefore, (2) is true: a priori physicalism is false.[\(4\)](#)

In support of (3) are considerations of the causal closure of the physical. Causal closure is the conjunction of two distinct theses. The first is a thesis about events; it is that for all physical events e , if there is an event e^* such that e^* causes e , then e^* is a physical event. The second is a thesis about properties; it is that if any property is causally efficacious in one physical event's causing another, that property is a physical property. Of course, neither these theses nor their conjunction implies that there are no irreducibly mental events or properties. What causal closure does plausibly imply however is that irreducibly mental events and properties can play no causal role in the production of physical events, and so, of behavior. But the thesis that the mental has no causal work to do in the production of behavior *just is* epiphenomenalism. By considerations of causal closure, therefore, (3) is true: if physicalism is false, epiphenomenalism is true.[\(5\)](#)

In support of (4) are considerations of evidence, and, in particular, considerations of what constitutes evidence for the existence and instantiation of qualia. In both ourselves and others, we come to know about qualia via systems of memory, introspection and perception. According to one natural approach to these systems, however, they provide evidence of aspects of the world only if those aspects stand in an appropriately direct causal link to us. But if epiphenomenalism is true, it is hard to see how qualia might stand in any such link, and thus it is hard to see our reason for saying there are qualia in the first place. By considerations of evidence, therefore, (4) is true: epiphenomenalism is false.[\(6\)](#)

Of course, the fact that each of (1-4) has arguments in its favor does not affect the possibility that, on reflection, one of them might turn out to be false. Indeed, interpreted in the simplest possible way, most recent contributions to the debate (and in fact most classical contributions) are arguments to the effect that, despite initial plausibility, one or more of the theses is false. Thus, a posteriori physicalists reject (1): according to them, it is not the case that the only way of justifying or explaining the supervenience thesis implied by physicalism is by making plausible an analysis of the mental into the physical. A priori physicalists reject (2): according to them, either physical knowledge can suffice for qualitative knowledge after all, or else the sense in which it cannot poses no threat to a priori physicalism. Interactionist dualists reject (3): according to them, physical closure is false, either because some physical events are caused by irreducibly mental events, or else because the properties which are efficacious in the production of some

physical events are irreducibly mental. And epiphenomenalists reject (4): according to them, one thing can be evidence for another even if no direct causal relation obtains between them.

Certainly the rational weight behind each of these options might in the end be such that we ought to endorse it. But it is important to see that rejecting one or more of its constituent theses is only one way of resolving the physicalism debate. Another possibility is to argue for some kind of ambiguity. If we can discern an ambiguity in (1-4) then it would seem possible to believe *all* the theses that make up the debate, rather than rejecting one of them.[\(7\)](#)

My aim in this paper is to suggest that it *is* plausible to discern an ambiguity in (1-4), and to defend a version of physicalism on that basis. There are, I think, two rather different conceptions of the physical, and hence of physicalism, at play in the physicalism debate. What I want to suggest is that if these two conceptions are distinguished, the apparent inconsistency of (1-4) disappears. More particularly, if the two conceptions are distinguished, it becomes clear that there is no *one* notion of the physical according to which it is plausible to believe both (2) and (3): the sense of 'physical' in which a priori physicalism is false is not the sense in which the rejection of physicalism inevitably leads to epiphenomenalism.

This possibility will of course recommend itself to those who find each of (1-4) attractive. But it is also of interest for another reason. Much of the recent discussion in philosophy of mind has had physicalists either denying that one can have propositional knowledge concerning qualia or else appealing to what Chalmers (1996) calls 'strong necessities'-necessities which are a posteriori but are totally unlike the sort of a posteriori necessities discussed by Kripke (1980). Perhaps it will turn out that one cannot have propositional knowledge of qualia, or that there are strong necessities, but it would be helpful if we had a way of defending physicalism that did not require taking a stand on such issues.[\(8\)](#)

My approach is as follows. In §2, I set out the two conceptions of the physical I will be interested in. In §§3 and 4, I suggest that distinguishing the two conceptions resolves the contemporary debate. In §5, I defend this suggestion against some objections, with a particular focus on the sort of physicalism one is left with if the suggestion is adopted—as we will see, the position that emerges is similar in important respects to Russell's neutral monism, and to modern versions of that view. Finally, in §6, I close the paper by providing some positive reasons for endorsing the strategy I propose.

2. According to the first conception of the physical—which I will call *the theory-based conception*—a physical property is a property which *either* is the sort of property that physical theory tells us about *or* else is a property which metaphysically (or logically) supervenes on the sort of property that physical theory tells us about. According to this conception, for example, if physical theory tells us about the property of having mass, then having mass is a physical property. Similarly, if physical theory tells us about the

property of being a rock-or, what is perhaps more likely, if the property of being a rock supervenes on properties which physical theory tell us about-then it too is a physical property. Let us say that any property which is physical by the lights of the theory-based conception is a *t-physical property*.⁽⁹⁾

According to the second conception of the physical-which I will call *the object-based conception*-a physical property is a property which *either* is the sort of property required by a complete account of the intrinsic nature of paradigmatic physical objects and their constituents *or* else is a property which metaphysically (or logically) supervenes on the sort of property required by a complete account of the intrinsic nature of paradigmatic physical objects and their constituents. ⁽¹⁰⁾ According to this conception, for example, if rocks, trees, planets and so on are paradigmatic physical objects, then the property of being a rock, tree or planet is a physical property. Similarly, if mass is a paradigmatic physical object-or, what is perhaps more likely, if the property of having mass is required in a complete account of the intrinsic nature of physical objects and their constituents-having mass is a physical property. Let us say that any property which is physical by the lights of the object-based conception is an *o-physical property*.

Since it is one of the more central and difficult notions that we have, many issues will arise whenever one is discussing *any* notion of the physical. However, the issue I want to focus on here is the sense in which the two conceptions I have described are distinct. ⁽¹¹⁾ Of course, it is obvious that they are distinct in some sense: one concerns theories, the other objects. But what is of interest to me is whether the class of properties characterized by both is co-extensive. I will argue that they are not co-extensive for the following reason: some o-physicals are not t-physical.

The point emerges most clearly if we have two theses before us. The first is that physical theory tells us only about the *dispositional* properties of physical objects and so does not tell us about the *categorical* properties, if any, that they have. A thesis of this sort has been held (at least implicitly) by many philosophers, but the following passage from Blackburn provides an effective illustration:

When we think of categorical grounds, we are apt to think of spatial configurations of things-hard, massy, shaped things, resisting penetration and displacement by others of their kind. But the categorical credentials of any item in this list are poor. Resistance is *par excellence* dispositional; extension is only of use, as Leibniz insisted, if there is some other property whose instancing defines the boundaries; hardness goes with resistance, and mass is knowable only by its dynamical effects. Turn up the magnification and we find things like an electrical charge at a point, or rather varying over a region, but the magnitude of a field at a region is known only through its effects on other things in spatial relations to that region. A region with charge is very different from a region without...It differs precisely in its dispositions or powers. But science finds only dispositions all the way down. (1992, pp. 62-3)

What Blackburn is saying here is that when we consider both the properties of physical objects we normally think of as primary qualities-such as resistance-and the physical properties we normally think of as being associated with modern physics-such as having a certain charge-we find that such properties are dispositional. His final remark suggests that something more general is true and that in scientific theory-Blackburn means, I think, *physical* scientific theory-one only "finds" dispositional properties. It seems reasonable to summarize this by saying that physical theory tells us only about dispositional properties.[\(12\)](#)

The second thesis we need to consider is that the dispositional properties of physical objects *do* require categorical grounds, i.e. for all dispositional properties, there must be a non-dispositional property, or non-dispositional properties, such that the instantiation of latter is metaphysically sufficient for the instantiation of the former. For example, if a vase is fragile, there must be a non-dispositional property, or non-dispositional properties, whose instantiation makes it the case that the vase is fragile; and if a chair is uncomfortable, there must be a non-dispositional property, or non-dispositional properties, whose instantiation makes it the case that the chair is uncomfortable. [\(13\)](#) It might perhaps be thought that there is some conflict between this thesis and the first. But there is certainly no *logical* inconsistency here. And indeed, many (but not all) philosophers who hold the first thesis also hold the second. A prominent example is D.M. Armstrong, who in *A Materialist Theory of the Mind*, holds that, since physical theory only characterizes the dispositional or relational nature of physical objects, it therefore "does not tell us"-as Armstrong (1968, p. 282) says- about their categorical or non-relational nature.

With these two theses in place it is easy to argue that the theory-based conception is distinct from the object-based conception. Suppose a physical object x has a dispositional physical property F . From the thesis that dispositional properties require categorical grounds, it follows that x (or its constituents) must also have a further non-dispositional property, which we may call G . But now let us ask: is G a physical property or not? If we are operating with the theory-based conception, it would seem that G is *not* physical. For from the thesis that physical theory tells us only about dispositions, it follows that t-physical properties are either dispositional or else supervene on dispositional properties. But neither is true in the case of G ; so G itself is not t-physical. On the other hand, if we are operating with the object-based conception, there is no reason at all to deny that G is a physical property. After all, G is-or at least could perfectly well be-the kind of property required in a complete account of the intrinsic nature of paradigmatic physical objects and their constituents. In sum, properties such as G -the properties which are the categorical grounds of the dispositional properties that physical theory tells us about-serve to show that the two conceptions of the physical are distinct. By the lights of the theory-based conception, G and its ilk are not physical; but by the lights of the object-based conception, G and its ilk are.

The chief complication with this argument-apart of course from the issues that we have already set aside (cf. fn 12 and 13)-derives from the fact that there are two different senses in which physical theory might fail to tell us about the categorical properties of physical objects. In the first sense, physical theory fails to tell

us about a property just in case no expression of the theory *refers* to that property. Now, in this sense, the argument we just considered will not go through. For even if dispositional properties do require categorical grounds, it is still perfectly possible that the expressions of physical theory might refer to those grounds *inter alia*. But then *G* and its ilk will count as t-physical as well as o-physical, and there is no reason for supposing that the two conceptions of the physical are distinct.

But there is also a second sense in which physical theory might fail to tell us about categorical grounds. In this sense, physical theory fails to tell us about categorical grounds just in case there might be two possible worlds *w* and *w** such that (i) they are exactly alike in terms of their distribution of dispositional properties-in both *w* and *w**, *x* has *F*; but (ii) they are different in terms of their categorical properties-in *w*, *x* has *G* but in *w**, *x* has a quite distinct categorical property *G**; and (iii) they are from the point of view of the theory epistemically indiscernible-the two worlds are (in Kripke's famous phrase) epistemically and qualitatively identical, though they might nevertheless be semantically different in the sense that in *w*, an expression of physical theory refers to *G* while in *w** the counterpart expression refers to *G**. It is this idea that Blackburn and Armstrong are appealing to when they (respectively) say that science finds dispositions all the way down, and that physical theory does not tell us about the categorical nature of physical objects. Similarly, it is this idea that is required by the argument we just considered that some o-physical properties are not t-physical.[\(14\)](#)

The right response to this complication is to acknowledge it and set it aside. Suppose there *are* two senses of what it is for a theory to tell us about a property. Then we will of course have *two* different versions of the theory-based conception of the physical. Moreover, only one of these versions is distinct from the object-based conception. On the other hand, the two different versions of the theory-based conception will *themselves* determine two different classes of properties: according to the version which employs the first sense of 'tells us about', the theory-based conception of the physical will acknowledge both dispositional and categorical properties as physical; according to the version which employs the second sense, the theory-based conception will acknowledge only dispositional properties as physical. And thus our basic point remains the same: there is distinction between two conceptions of the physical and these two conceptions determine two classes of physical properties, the first of which is limited to the dispositional properties of physical objects, the second of which includes both the dispositional and categorical properties of physical objects. More generally, one could run the following discussion in one of two ways: either one could operate with the two versions of the theory-based conception, or else one could operate with the object-based conception and the version of the theory-based conception which employs the second sense of 'tells us about'. I will here adopt the latter course-and thus I will continue to contrast o-physical properties and t-physical properties-but it is important to notice that the former course is also available.

3. So far, I have introduced two conceptions of the physical, and argued that they characterize distinct classes of properties. I turn now to the suggestion that distinguishing the two conceptions provides a way of believing all the constituent theses in the physicalism debate.

The first thing to say is that the issue of the interpretation of the notion of the physical does not seem to affect the reasons for believing either thesis (1) or thesis (4). The arguments that support (1)-viz., the claim that if physicalism is true a priori physicalism is—are arguments of a very general nature about supervenience, and whether or not the defense of supervenience will in the end involve an a priori entailment of the mental by the physical. These arguments are of course highly controversial. But, however things turn out with these arguments, it seems plain that the issues they raise are orthogonal to the issues having to do with the conceptions of the physical.

The same thing applies to the arguments one might adduce in support of (4)-viz., the claim that epiphenomenalism is false. These are arguments to the conclusion that if epiphenomenalism is true, the systems by which we gather evidence of qualia break down. But once again, while these arguments are controversial, I think we can set them aside. However things turn out with these arguments, it seems plain that the issues they raise are irrelevant to the issues with which we are concerned.

However, while the issue of interpretation does not affect the reasons for (1) and (4), it *does* affect the reasons for (2) and (3).

Let us consider first (2), the claim that a priori physicalism is false. Given our distinction between the theory-based and the object-based conceptions, we can distinguish two interpretations of (2). According to the first, (2) asserts that it is not the case that qualia supervene on t-physical properties *and* that t-physical truths a priori entail qualitative truths. We might express this by rendering (2) as:

(2-t) A priori t-physicalism is false.

According to the second interpretation, (2) asserts that it is not the case that qualia supervene on o-physical properties *and* that o-physical truths a priori entail qualitative truths. We might express this by rendering (2) as:

(2-o) A priori o-physicalism is false.

It is obvious that (2-t) and (2-o) are, given our assumptions, distinct. After all, if there are properties which are o-physical but not t-physical, it follows that even if t-physicalism is false, o-physicalism might still be true.[\(15\)](#)

Now, earlier I said that what supports (2)-and so counts against a priori physicalism-are considerations of epistemic distinctness. Given that (2) can be disambiguated into (2-t) and (2-o), what we need to ask is whether these considerations support both disambiguations. I think it can be argued, however, that while the considerations of epistemic distinctness do support (2-t), they do not likewise support (2-o). And this means that we can maintain a priori o-physicalism in the face of the considerations of epistemic distinctness even while we cannot maintain a priori t-physicalism.

While there are a number of different kinds of consideration of epistemic distinctness, we will focus here on what is perhaps the clearest and most notorious of them, Frank Jackson's (1982) knowledge argument.⁽¹⁶⁾ As is extremely well-known, this argument asks us to imagine Mary, a famous neuroscientist confined to a black and white room. Mary is forced to learn about the world via black and white television and computers. However, despite these hardships Mary learns (and therefore knows) all that physical theory can teach her. Now, if a priori physicalism were true, it is plausible to suppose that Mary knows everything about the world. And yet-and here is Jackson's point-it seems she does not know everything. For, upon being released into the world of color, it will become obvious that, inside her room, she did not know what it is like for both herself and others to see colors-that is, she did not know about the qualia instantiated by particular experiences of seeing colors. Following Jackson (1986), we may summarize the argument as follows:

(5) Mary (before her release) knows everything physical there is to know about other people.

(6) Mary (before her release) does not know everything there is to know about other people (because she learns something about them on being released).

Therefore,

(7) There are truths about other people (and herself) that escape the physicalist story.

And of course, the truth of (7) entails (2), the thesis that a priori physicalism is false; for, if a priori physicalism were true, there would be no truths about anybody (or anything) that escape the physicalist story. For our purposes, therefore, it is reasonable to by-pass (7) and interpret the knowledge argument simply as urging that (5) and (6) entail (2).

Given our distinction between the theory-based and the object-based conceptions of the physical, however, it is clear that the first premise of this argument is subject to interpretation in either of two ways. We might express these as (5-t) and (5-o):

(5-t) Mary (before her release) knows everything t-physical there is to know about other people.

(5-o) Mary (before her release) knows everything o-physical there is to know about other people.

Moreover, it is clear that (5-t) and (5-o) differ both in plausibility and in what conclusions they support. (5-t) is certainly very plausible given the story of Mary. After all, if Mary knows all of physical theory, and if physical theory tells us about t-physical properties, she presumably will know everything t-physical there is to know about other people. But the trouble here is that, when combined with (6), (5-t) only yields the falsity of a priori t-physicalism, not a priori o-physicalism, i.e., it only yields (2-t), not (2-o). If some o-physical properties are not t-physical, from the fact that Mary knows everything t-physical about the world, it does not follow that she knows everything o-physical, and it therefore does not follow that when she learns something about the world she has learnt something non-physical. The truth of (5-t) and (6), then, leaves us free to endorse a priori o-physicalism, and therefore free to reject (2-o).

But what if we operate with (5-o) rather than (5-t)? It is obvious that the conjunction of (5-o) and (6), unlike the conjunction of (5-t) and (6), *does* entail the falsity of a priori o-physicalism, i.e., does entail (2-o). But the trouble now is that the story of Mary gives us no reason at all to endorse (5-o). The reason for this is that, on the construal that we have adopted, physical theory *will not tell us* about certain of the o-physical properties. After all, if the physical theory does not tell us about the categorical properties of physical objects, and if some o-physical properties are categorical, it would seem that, no matter how much physical theory one knows, one will still not know about certain o-physical properties. But then no matter how much physical theory *Mary* knows, she will still not know about certain of the o-physical properties. Nothing she knows, therefore, rules out the possibility that the truths about herself and others are entirely o-physical. More generally, while (5-o) and (6) certainly provide premises from which one might validly argue to the falsehood of a priori o-physicalism, the argument is a failure nevertheless, because we have no reason to believe its first premise.

In sum, the distinction between the theory-based conception and the object-based conception provides a way in which we might defeat the knowledge argument and at the same time concede the central intuition that motivates it. [\(17\)](#)In consequence, we can both accept and deny thesis (2). On one interpretation, (2) is equivalent to (2-t), and we have seen that the knowledge argument gives us very good reason to grant this. On the other interpretation, (2) is equivalent to (2-o), and we have seen that the knowledge argument gives us no good reason to grant this. To that extent then, we can with good conscience accept (2) if it is interpreted as (2-t), and reject it if it is interpreted as (2-o).

4. The suggestion that one might reject (2) on one interpretation and accept it on another is an important step in resolving the puzzle posed by the inconsistency of (1-4). But this alone does not complete the resolution. For I have conceded that (2) is very plausible on the theory-based conception of the physical; that is, that (2-t) is very plausible. And I have also suggested that the truth of (1) and (4) remains untouched by any issue about how to interpret the physical. But this means that if (3)-viz., the claim that if physicalism is false, epiphenomenalism is true-is plausible on the theory-based conception, we are back where we started. To complete the resolution of the puzzle, therefore, it needs to be argued that (3) is implausible if interpreted from the standpoint of the theory-based conception.

However, (3) *is* implausible if interpreted from the standpoint of the theory-based conception.

Let us first distinguish the two interpretations of (3) just as we did for (2). On the first interpretation, (3) asserts that if qualia do not supervene on t-physical properties, the inevitable result is epiphenomenalism. We might express this by rendering (3) as:

(3-t) If t-physicalism is false, epiphenomenalism is true.

On the second interpretation, (3) asserts that if qualia do not supervene on o-physical properties, the inevitable result is epiphenomenalism. We might express this by rendering (3) as:

(3-o) If o-physicalism is false, epiphenomenalism is true.

Once again, it is clear that (3-t) and (3-o) are, given our assumptions, distinct. If there are o-physical properties which are not t-physical, it is clear that we might have grounds for rejecting (3-t) but no grounds at all for rejecting (3-o).

Now, earlier I said that the considerations in favor of (3) are considerations of causal closure. And we also saw that causal closure is the conjunction of two theses, one about events, the other about properties. We might summarize these theses as (8) and (9):

(8) For all physical events e , if there is an event e^* such that e^* causes e , then e^* is a physical event.

(9) For all physical events e and e^* , if there is a property F such that F is causally efficacious in e 's causing e^* , then F is a physical property.

Now, (8) can from our point of view be set aside. The reason is that the conceptions of the physical with which we are operating apply in the first instance to *properties*, and do not obviously extend to items of

other ontological categories. More particularly, neither conception as it stands says anything about what it is to be a physical event. But this means that the conceptions are silent on the plausibility of (8).

On the other hand, the conceptions are not silent on the plausibility of (9). If we interpret (9) in accordance with the theory-based conception, it asserts that the only properties that are causally efficacious are *t-physical* properties, a thesis we might call (9-t):

(9-t) For all physical events *e* and *e**, if there is a property *F* such that *F* is causally efficacious in *e*'s causing *e**, then *F* is a *t-physical* property.

If we interpret (9) in accordance with the object-based conception, by contrast, it asserts that the only properties that are causally efficacious are *o-physical* properties, a thesis we might call (9-o):

(9-o) For all physical events *e* and *e**, if there is a property *F* such that *F* is causally efficacious in *e*'s causing *e**, then *F* is an *o-physical* property.

And once again, it is clear that (9-t) and (9-o) are, given our assumptions, distinct. The first implies that only *t-physical* properties are causally efficacious, while the second allows that some causally efficacious properties are not *t-physical*, i.e., those properties which are *o-physical* but not *t-physical*.

Now, I think we should concede that *if* (9-t) is true, we have a very good argument for (3-t). Similarly, we should concede that *if* (9-o) is true, we have a very good argument for (3-o). For consider: if (9-t) is true, the only causally efficacious properties in the production of physical events, and therefore of behavioral events, are *t-physical*; but if *t-physicalism* is false, qualia are not *t-physical*. It follows that qualia are not efficacious in the production of behavior. *Mutatis mutandis* for (9-o): if (9-o) is true, the only causally efficacious properties in the production of physical events, and therefore of behavioral events, are *o-physical*; but if *o-physicalism* is false, qualia are not *o-physical*. It again follows that qualia are not efficacious in the production of behavior.

On the other hand, there is an important difference between (9-t) and (9-o), and this is that it is far from obvious that (9-t) is *true*. For let us consider more directly what (9-t) says. (9-t) says that in the realm of physical causation, the only causally efficacious properties are *t-physical*. In general, however, this seems to be quite mistaken. The reason is that it is very implausible to suppose that the efficacy of dispositional properties has nothing whatsoever to do with the efficacy of their categorical grounds.

To illustrate this point, consider Ned Block's famous example of the bull and the bull-fighter. Block writes:

Consider the bull-fighter's cape. The myth (which we will accept, ignoring the inconvenient color-blindness of bulls) is that its red color provoked the bull, i.e., redness

is causally relevant to the bull's anger. The cape also has the second order property of being provocative, of having some property or other that provokes the bull, of having some property or other that is causally relevant to the bull's anger. But does the provocativeness of the cape provoke the bull? Is the provocativeness causally relevant to the bull's anger? It would seem not. The bull is too stupid for that. The provocativeness of the cape might provoke the ASPCA, but not the bull.(1990, p.155)

The moral that Block wants to draw from this example is the contentious one that provocativeness and similar properties are in a large class of cases not causally efficacious ("causally relevant") while properties such as redness are. For our purposes, it is sufficient to draw the less contentious moral that while dispositional properties are causally efficacious-for of course provocativeness is a dispositional property- they are only efficacious if their categorical grounds are-for of course redness is (or anyway may be taken to be for the purposes of the example) the categorical ground of provocativeness. As Mark Johnston puts it in a related context, if dispositional properties are efficacious, they are so "at one remove and by courtesy" (1992; p 235). But this means that (9-t) is false. (9-t) tells us that the only properties which are causally efficacious are dispositional, i.e., properties similar to provocativeness. But it seems clear that, if dispositional properties *are* causally efficacious, so too are their categorical grounds. And, if categorical grounds are, as we are assuming, non-dispositional, then, contrary to (9-t), dispositional properties are not the only properties that are causally efficacious.

On the other hand, examples such as Block's do nothing whatsoever to undermine (9-o). (9-o), after all, does not restrict itself to dispositional properties, and so is not subject to the criticism that (9-t) is. More generally, if we interpret (9) in accordance with the theory-based conception, we can reasonably regard it as false and therefore as providing no support for (3). On the other hand, if we interpret (9) in accordance with the object-based conception, we can reasonably regard it as true and therefore as providing support for (3). In summary, (9-o) provides support for (3-o), but (9-t) provides no support for (3-t).

How does this bear on the resolution of (1-4)? Well, earlier we saw that thesis (2) has two versions-viz., the version expressed by (2-t) and the version expressed by (2-o)-and that only the *first* of these receives support from considerations of epistemic distinctness; that is, considerations of epistemic distinctness leave us free to reject (2-o). What we have just seen is that thesis (3) also has two versions-viz., the version expressed by (3-t) and the version expressed by (3-o)-and that only the *second* of these receives support from considerations of closure; that is, considerations of closure leave us free to reject (3-t). But this suggests that the predicament which in fact confronts us not quite the predicament we originally imagined ourselves to be in.

Originally, we imagined ourselves to be confronted with four inconsistent theses (1-4) each of which we had powerful reason to believe:

- (1) If physicalism is true, a priori physicalism is true.
- (2) A priori physicalism is false
- (3) If physicalism is false, epiphenomenalism is true.
- (4) Epiphenomenalism is false.

If we distinguish between the two conceptions of the physical, however, it is plain that (1-4) fail to articulate *precisely* the theses we have reason to believe. What we in fact have reason to believe is not (1-4) but rather:

- (1) If physicalism is true, a priori physicalism is true.
- (2-t) A priori t-physicalism is false.
- (3-o) If o-physicalism is false, epiphenomenalism is true.
- (4) Epiphenomenalism is false.

But this second tetrad is importantly different from first. The first is inconsistent, the second is not. As a consequence, if you find the considerations of supervenience, epistemic distinctness, closure and evidence compelling, you may believe the second tetrad without fear of contradiction. And our original puzzle is solved.

5. My strategy so far has been to distinguish two conceptions of the physical, and to suggest that, if these conceptions can be kept apart, the appearance of inconsistency of (1-4) is resolved. If we adopt the object-based conception, we can reasonably regard (2) as true and (3) as false. But if we adopt the theory-based conception, we can reasonably regard (3) as true and (2) as false. I turn now to some objections to this strategy.

Broadly speaking, there are two classes of objection to consider. The first class raises questions about the assumptions I required in order to argue that, properly understood, (1-4) do not present a contradiction: the thesis that dispositions require categorical grounds, and the thesis that physical theory tells us only about dispositional properties. The second class raises questions about the outline of the position that one is left with if one pursues our strategy. Obviously, it is no good to resolve the inconsistency of (1-4) only to be forced into an even more unpalatable position. I have already noted that I will set aside the first class of objection, and so here I will concentrate on the second class.

What sort of position is one left with if one accepts (1), (2-t), (3-o), (4) and rejects (2-o) and (3-t)? The answer is that one is left with a view that bears a close resemblance to Russell's in *The Analysis of Matter*, and even closer resemblance to the view discussed by contemporary defenders of a Russell-inspired physicalism such as Maxwell (1978) and Lockwood (1989, 1992).[\(18\)](#)

The broad contours of the position may be brought out by considering the following analogy. Imagine a mosaic constituted by two basic shapes, triangles and pieces of pie, as well as a large number of shapes obtained by a transparent combination of these: squares, half-moons, circles, rhombuses etc. Imagine also that our access to the mosaic is limited to two shape-detecting systems: the first scans the mosaic and detects triangles; the second scans it and detects circles. For one reason or another we spontaneously assume that the triangle-detector tells us everything about the nature of the mosaic—we become trianglists, i.e. those who believe that triangles are the fundamental shape and that all other shapes supervene. The problem of the circle then stares us in the face: the circle-detector tells us the mosaic contains circles, but there is apparently no place for circles in a mosaic totally constituted by triangles. Different people respond to the problem in different ways: some say the circle-detecting system leads us astray, and that properly understood it provides no propositional knowledge of circles; others declare circles a posteriori identical with triangles; still others decide that circles are irreducible, and postulate contingent laws linking them and triangles. Of course all these responses are mistaken, and moreover they make a common mistake. The mistake is that the triangle-detector does not tell us everything about the mosaic: it is selective and only tells us about triangles when in addition there are pieces of pie. Of course the pieces of pie are not themselves circles. But in combination they may constitute circles. When God created the mosaic, all he had to do was to create triangles and pieces of pie, and arrange them in just the way he wanted; in doing so, he created everything else including circles.

As it is with triangles, pieces of pie, and circles so (largely) it is with t-physical, o-physical and qualia, according to the Russell-inspired view. Physical theory tells us about the physical world, and introspection tells us about qualia. For one reason or another we spontaneously assume that physical theory tells us everything about the nature of the world—we become t-physicalists, i.e., those who believe that t-physical properties, properties that physical theory tells us about, are fundamental and all other properties supervene. The problem of qualia then stares us in the face: introspection tells us there are qualia, but there is apparently no place for qualia in a world totally constituted by t-physical properties. Different people respond to the problem in different ways: some say introspection leads us astray, and that properly understood it provides no propositional knowledge of qualia; others declare qualia a posteriori identical with t-physical properties; still others decide that qualia are irreducible, and postulate contingent laws linking them and t-physical properties. Of course all these responses are mistaken, and moreover they make a common mistake. The mistake is that physical theory does not tell us everything about the physical world: it is selective and only tell us about dispositional t-physical properties when in addition there are categorical o-physical properties. Of course, the categorical o-physical properties are not themselves qualia. But in combination—perhaps also in combination with the t-physicals—they may constitute qualia. When God created the world, all he had to do was create the fundamental physical properties—o-physical and t-physical—and arrange them in just the way he wanted; in doing so, he created everything else including qualia.

To put things less picturesquely, the Russell-inspired position as I will understand it here has two parts. The first part-the physicalist part-is that qualia supervene not on the class of properties that physical theory tells us about-the t-physical properties, as I have called them-but on a larger class that includes both the t-physicals and categorical bases of such properties- o-physical properties, as I have called them. The second part-the a priori part-is that mental truths (in particular, qualitative truths) are a priori entailed by physical truths. In considering the following objections, then, I will be interested in the nature and plausibility of this Russell-inspired view, which I will continue to call a priori o-physicalism, or, more simply, o-physicalism.[\(19\)](#)

Objection#1: Is this really physicalism?

The first objection is whether o-physicalism is really physicalism. An objection of this sort is in effect mentioned by Chalmers. In the course of providing an otherwise extremely sympathetic discussion of the view I have been discussing, Chalmers says that "there is a sense in which this view can be seen as a monism rather than a dualism, but it is not a materialist monism" (1996; p.155). [\(20\)](#) Rather, he says, the view is either panpsychism or neutral monism, depending on how it is developed. If correct, this seems to constitute a major reason to doubt our claim to have found a resolution of the debate over physicalism: for we would have resolved the debate simply by giving up physicalism.

The general response to this objection is that it misses the distinction between the theory-based and the object-based conception of the physical. If one operates *only* with the theory-based conception-as Chalmers in effect does (1996; p 33)-then any property which is not a t-physical property will need to be excluded from the class of physical properties, and thus must be classed as either a mental property or a neutral property, i.e. a property which is neither mental nor physical. But if one accepts that there are more physical properties than the t-physicals, the landscape looks somewhat different.

To illustrate this, let us examine first the question of whether o-physicalism is a version of neutral monism. The crucial consideration here is that neutral monism as usually defined *presupposes* the theory-based conception of the physical: neutral properties are properties which are neither mental nor t-physical; and neutral monism is the thesis-to put it in modern terms-that mental and t-physical properties metaphysically supervene on neutral properties. Now, if this is what neutral monism is, there is certainly some plausibility in supposing that o-physicalism is a version of neutral monism. Since some o-physical properties are neither mental nor t-physical, those o-physical properties will be neutral in this sense. On the other hand, it is less clear that this feature of the view compromises its status as physicalism. If one accepts that there are physical properties which are not t-physical, then one way to hold this first version of neutral monism is to

adopt o-physicalism. So the charge that o-physicalism is not a version of physicalism rests on a failure to distinguish the two conceptions of the physical.

One might be tempted at this point to reject the object-based conception of the physical and to operate only with the theory-based conception. Of course, if one operates only with the theory-based conception, one may draw a clear distinction between neutral monism and physicalism. However, on the assumption that dispositions require categorical grounds, this option is not available to physicalists. For suppose you are a physicalist who believes that dispositional properties require categorical grounds. If you in addition believe that only dispositional properties are physical, it follows that for every dispositional physical property that is instantiated a non-physical categorical property is instantiated also. But then physicalism is false-and for reasons that have nothing to do with philosophy of mind. In sum, if a physicalist wants a version of physicalism according to which the doctrine can be true compatibly with certain metaphysical assumptions about dispositions, it is necessary that he or she operates with a conception of the physical broader than simply the theory-based conception. In the context of our argument, this means that physicalists must adopt the object-based conception.[\(21\)](#)

So there seems little reason to be concerned that o-physicalism is a form of neutral monism-but what then of panpsychism? According to panpsychism, the categorical properties which underlie dispositional t-physical properties are in every case qualia. It follows from panpsychism therefore that all the physical objects of our acquaintance-computers, trees, planets etc-all instantiate qualia just as I do. Now, while this view is somewhat startling, it can be given a motivation. Part of its motivation is once again the idea that (dispositional) t-physical properties exhaust the class of physical properties. But in part also its motivation is that we seem to glean the concept of a categorical property from our concepts of qualia-as Blackburn (1992; p. 65) puts it, "categoricity comes with the subjective view". Putting these two ideas together, it becomes tempting to suppose that, in general, physical properties require categorical bases, and that these categorical bases are simply qualia. The result is panpsychism.

Now of course the first part of this motivation will be rejected by someone who accepts the object-based conception of the physical. But the second part of the motivation is also open to objection. For even if one derives one's concept of a categorical property from one's concept of qualia-and this is something we can remain neutral on-it does not follow that all categorical properties are qualitative properties. In Kripke's (1982; p.118) example, even if one derives one's concept of a duck from the ducks in Central Park, it does not follow that the concept so derived does not apply to ducks not in Central Park. And this means that one can at least imagine a range of categorical properties which are both physical and non-qualitative: these are the properties which make up the categorical nature of physical objects, the properties we have called o-physical. If o-physicalism is right, then some of these properties will in combination be the supervenience base for qualitative properties. Unless there is something incoherent in the very idea of such a class of properties, we have no reason to suppose that o-physicalism will collapse into panpsychism.

Objection #2 Concepts unattainable?

The second objection I will consider concerns the a priori part of o-physicalism: the thesis that the qualitative truths are a priori entailed by physical truths.

It needs to be appreciated that in committing itself to such physical truths o-physicalism is committing itself to a class of truths which cannot be expressed in a language we currently understand, or, to put it in another idiom, cannot be formulated using concepts we currently possess. For consider: what could the relevant concepts be? They cannot be (what we might call) *t-concepts*, i.e. concepts which tell us about t-physical properties-for in our discussion of the knowledge argument we in effect conceded that such t-concepts are not a priori equivalent to concepts of qualia. And nor can they be concepts of qualia-for the position is precisely that the concepts in question are physical not qualitative. So o-physicalism requires the postulation of a third class of concept, which we might call *o-concepts*: these are concepts which tell us about the categorical o-physical properties of physical objects, just as qualitative concepts tell us about qualia, and t- concepts tell us about the dispositional t-physical properties of physical objects.

Now, the mere fact that o-physicalism is committed to a class of concepts which we do not currently possess is not by itself an objection against the view. For one thing, any sober assessment of our epistemic achievements must always acknowledge the possibility that our conceptual repertoire is in various respects limited. Moreover, the concepts at issue here are concepts which for most practical purposes do not matter, and so the fact that we lack them is not going to disrupt our usual epistemic engagement with the world. Those philosophers who hold that it is dispositions 'all the way down', and thus that there are no categorical properties, are in my view mistaken. But the grain of truth in their position is that the distinction between categorical and dispositional does not seem to matter much to the main business of science.

Nevertheless, one might try to develop this feature of our account-that it requires a new class of concept-into an objection against it in a number of ways. One suggestion is that, not only do we not possess o-concepts, no possible being *could* possess them. However, while it is difficult to argue against this suggestion it is also difficult to see the motivation for it. In general we *know* that one can possess concepts relevantly like o-concepts. As Blackburn (1992) points out, concepts of qualia are precisely such concepts, at least in the sense that these are concepts of non-dispositional properties. Why then, on an analogy with such concepts, can we not imagine a being who possessed concepts of the categorical properties of physical objects? The claim that no possible being could possess these concepts might have some force if we did not possess concepts of qualia. But the trouble is that we *do* possess such concepts.

Alternatively, one might argue that the idea of a class of o-concepts is incoherent, on the ground, for example, that it is simply incoherent to suppose that physical truths could a priori entail qualitative truths.

However, the problem with this suggestion is that we have at present no argument for it (even if such an argument could be developed in the future). One of the lessons of the knowledge argument is that t-concepts are not a priori equivalent to concepts of qualia. But it does not follow from this that o-concepts are not so equivalent. Again: what the knowledge argument shows is that a priori t-physicalism is false. But it does not show that a priori o-physicalism is false.

Of course, it remains an interesting and open question whether the contingent facts about our history or environment or psychology will allow us to develop the concepts at issue. One possibility is that we could develop such concepts, perhaps by undergoing an intellectual or scientific revolution akin to the revolutions of the 17th century. If that were true, our position would be similar to that represented by Nagel's (1974) famous presocratic philosopher who does not understand the identification of matter and energy. Another possibility is that our contingent psychological nature is such that we will never attain the o-concepts. If that were true, our position becomes a species of the position defended by McGinn (1989; see also Jackson 1982)-though we have arrived at the position from another direction. On this view, there are a class of concepts which would tell us about the categorical properties of physical objects, and which are required to formulate the physical truths which a priori entail the qualitative truths. But our contingent psychological nature is such that we shall never develop such concepts: we can develop concepts of the dispositional properties of physical objects; and we can develop concepts of qualia, but we cannot develop concepts of the categorical properties. Hence we are epistemically bounded with respect to what McGinn thinks of as 'the solution to the mind-body problem'.

It is not important for our purposes to decide between these possibilities. What is more important is to see the o-concepts as a kind of epistemic ideal. In general, we can imagine a kind of inquiry which would tell us about the categorical properties of physical objects-categorical inquiry, as we might call it. What our position predicts is that in order to have an a priori physicalist theory of qualia and their place in the world-one which provides an a priori analysis of the nature of the world in terms of a relative small list of physical concepts-one would need to complete the categorical inquiry. Whether we will ever be able to achieve this aim is an open question, but the goal remains as something to which we can aspire.

One might object that an account of the metaphysics of mind should not postulate epistemic goals and then offer no constructive suggestions about how that goal is to be reached. I agree that this is a drawback, and something that needs to be taken into consideration when one compares the present proposal with others. But it is possible also to take a somewhat more positive view of the situation. For one thing, the issue of epistemic goals needs to be sharply distinguished from the issue of the truth of physicalism. We are no longer debating whether qualia supervene on the physical or not, we are debating whether we ourselves can attain a certain epistemic goal, viz., come to possess a class of physical concepts from which the concepts of qualia would a priori follow. For another, a pleasing aspect of our position is that it allows us to

characterize the epistemic problem that faces us with relative precision. All that remains to be seen is if we can solve the problem

Objection #3: The Grain Problem

Finally, I consider the grain problem. This problem was originally raised by Sellars as a problem for physicalism in general, but many who have discussed or defended the Russell-inspired view have thought it the central problem that the doctrine faces- "certainly the hardest problem for any sort of Russellian View" (Chalmers 1996: 307). The reason the problem is pressing is because it threatens to show that, in principle, a (potentially huge and extremely complicated) collection of o-physical properties cannot constitute qualia. If this is right, regardless of its status as a version of physicalism, and regardless of the nature of the epistemic goals it envisions, our position is false. So no discussion of the Russell-inspired view can fail to address the grain problem.

While there are a number of statements of the grain problem in the literature, I will concentrate on Maxwell's. He writes:

The objection asks, for example, how it is that the occurrence of a smooth continuous expanse of red in our visual experience can be identical with a brain process that must, it would seem, involve particulate, discontinuous affairs such as transfers of or interactions among large numbers of electrons or the like. Surely being smooth is a structural property, and being particulate or discontinuous is also a structural property, one moreover that is incompatible with being smooth and continuous. This strongly, suggests, the objection continues, that at least some mental events exemplify structural properties that not exemplified by any brain event...It follows that the mental event and the brain event do not share all of their (structural) properties, and thus, the objector concludes, they cannot be identical (1978, p.398).

Here Maxwell is concentrating on identity, and thus one might be tempted to argue that his objection does not threaten o-physicalism. For of course our claim is that qualia *supervene* on o-physical properties, not that they are identical with them. However, I don't think this reply is going to carry much weight with a proponent of the grain objection. For the question can be stated without appealing to identity, as follows: how could a potentially non-continuous and unsmooth myriad of o-physical properties combine together to entail a smooth continuous expanse of red? Unless we can answer this more general question, the grain problem remains unsolved.

If the grain problem cannot be answered by distinguishing between identity and supervenience, how *can* it be answered? My own view [\(22\)](#) is that the answer emerges when we focus on what precisely it is in Maxwell's example that is supposed to be smooth and continuous. It seems plausible to say that it is the

expanse that is smooth and continuous, and also that the expanse is something that we represent in visual experience, i.e., Maxwell's example is an example in which we are having an experience which represents an expanse as being smooth and continuous. But of course, it does not follow from this that the experience *itself* is smooth and continuous. Consider: an experience of red represents something as being red, but it itself is not red. So the answer to the grain problem is that it gets the phenomenology wrong and mislocates the absence of grain: absence of grain is not a feature of experiences, but a feature of something that experiences represent.

One might object that it is obvious in introspection that one's experience of a red expanse is *itself* smooth and continuous and not simply that it represents an expanse as being so. However, this is not obvious at all. As many philosophers have emphasized, many acts or states of experiencing seem in a certain respect "diaphanous" to introspection: introspection reveals the intentional objects of experiences to us, but not the experiences themselves. As Moore put it: "when we try to introspect the sensation of blue, all we can see is the blue: the other element is as if it were diaphanous" (1922; p.25). To apply this to the case at hand, if we try to introspect the experience of a smooth and continuous expanse, all we see is the smoothness and continuity of the expanse: the other element-i.e., the experience itself-is as if it were diaphanous. If this account of the phenomenology of introspection is correct, it is by no means obvious that introspection reveals experiences to be smooth at all. What it reveals rather is that we often have experiences which represent things as being smooth and continuous. But that is a different matter.

Alternatively, one might object that to appeal to Moore's point about diaphanousness, and perhaps also to the difference between properties of experiences and properties represented by experiences, is to undercut the idea that there are qualia in the first place. However, while some philosophers-e.g., Harman 1990-have argued in this way, there is no certainly necessary connection here. As Shoemaker (1994) argues, it is perfectly possible to develop a theory of experience which honours the phenomenological point that Moore was making, and at the same time postulates qualia. On Shoemaker's preferred version of such a theory, an experience of a red expanse represents the expanse as having two properties-the property of being red (which Shoemaker thinks of as some physical property) and what he calls a *phenomenal property*, the property of causing an associated r-qualia, where an r-qualia is the type of qualia typically produced by red things. One way (not the only way) of developing Shoemaker's theory to handle the grain problem would be to argue that the experience of a smooth expanse represents the expanse as having two properties-smoothness (which we might think of as some physical property of a surface) and the phenomenal property of causing an associated s-qualia, where an s-qualia is the type of qualia typically produced by smooth things. Since the s-qualia is not itself smooth-just as the r-qualia is not itself red-it will not be the case that the experience or the qualia have structural properties that the o-physical properties lack; at any rate, this is something that might well be denied by a proponent of o-physicalism. But to deny this is to deny the central premise of the grain problem.

Of course, providing this sort of answer to the grain problem makes no progress at all on what one might suspect is really lying behind the grain problem, viz., whether we can *in fact* articulate a body of physical truths which a priori entail truths about r-qualia (or a s-qualia, if such there be). I have already admitted that our account makes no progress on this question beyond formulating in general terms what an answer to it would be like. The grain problem shows how hard this question is to answer, but it does not show that it is impossible to answer in principle.

6. I began by suggesting that the contemporary debate in philosophy of mind over physicalism might be viewed as involving an inconsistent tetrad of theses. Then I suggested that one plausible way of resolving the debate is by distinguishing between two conceptions of the physical, the object-based conception and the theory-based conception. What I have just done is defend the resulting version of physicalism against some objections.

Of course, even if our argument is right, it has only been shown that this is one way of resolving the debate. It has not been shown that this is the best or the only way. In order to provide anything approaching a full-dress defense of the proposal then, one would need to consider all the other attempts at resolving the puzzle presented by (1-4), and to demonstrate that these attempts are unsatisfactory, or at any rate less satisfactory than the proposal I have been considering. Obviously, there is no chance of my doing anything of the sort here. What I will do instead, however, is close the paper by making three points which are designed to show the *prima facie* plausibility of a proposal along the lines we have been considering.

The first point simply reiterates something from our discussion of physicalism and neutral monism. If one operates *only* with the theory-based conception of the physical, one ends up with a version of physicalism (and physical closure-cf fn 19) that, given certain metaphysical assumptions about dispositions and their categorical bases, is false. One is thus forced to introduce both the object-based conception of the physical and o-physicalism in order that one has a viable version of these theses at all. Hence if you want to be a physicalist-and if you accept those metaphysical assumptions-you had better be an o-physicalist.

The second point concerns the problem of other minds. Whether or not we are sceptics about other minds, we should all agree with the modal premise that sets the problem up. A version of that premise is:

(10) No amount of information of the sort presented in visual and sensory experience concerning another person *P* by itself entails that *P*'s experiences instantiate qualia.

Now, one sort of criticism of traditional formulations of physicalism-made, e.g., by Nagel (1970)-is that, on those formulations, the truth of physicalism makes it hard to see why the problem of other minds is a problem in the first place. I take this to mean that it is hard to see how physicalism and (10) could be true

together: on the assumption that physical information can in principle be presented as visual and sensory information (at least if the latter is enriched with causal and counterfactual notions), physicalism will entail that (10) is false. When taken as a criticism of t-physicalism, this objection seems to have some force. After all, when enriched with causal and counterfactual notions, the sort of information presented in visual and sensory experience might well amount to t-physical information; but if qualitative truths a priori follow from t-physical information, then (10) would be false. On the other hand, when taken as a criticism of o-physicalism, this criticism has no force. For the object-conception of the physical precisely allows two situations to be epistemically indiscernible but physically distinct: hence o-physicalism allows (10) to be true. Since the problem of other minds is a hard philosophical problem, and not simply an empirical concern, we have reason to prefer o-physicalism to t-physicalism.

The final point turns on a distinction between two kinds of strategies. A *rejectionist* strategy (as I will call it) is a strategy according to which we reject one or more of the theses (1-4) with which we began. An *accommodationist* strategy (as I will call it) is a strategy according to which we attempt to show, despite appearances, that (1-4) are jointly consistent. Now, other things being equal, it would seem that any accommodationist strategy is to be preferred over any rejectionist strategy. After all, as we saw in §1, it is not difficult to articulate reasons for believing each of (1-4). The disadvantage common to all rejectionist strategies is that they involve giving up something that we believe. As we saw at the outset, however, most contributions to the contemporary debate are (as usually interpreted) rejectionist: a posteriori physicalists reject (1); a priori physicalists reject (2); interactionist dualists reject (3); and epiphenomenalists reject (4). On the other hand, the strategy that we have been considering is obviously accommodationist-in fact it is the only accommodationist proposal currently on the table-and this gives us a strong prima facie reason to prefer it to any rejectionist strategy.[\(23\)](#)

Australian National University/ University of Colorado at Boulder

References:

- Armstrong, D.M. 1961. *Perception and the Physical World*. London: Routledge.
- Armstrong, D.M. 1968, p. 282. *A Materialist Theory of the Mind*. London: Routledge
- Blackburn, S. 1992. 'Filling in Space', *Analysis*, pp.60-65.
- Block, N. 1980. 'Troubles with Functionalism'. In Block, N. 1980 (ed) *Readings in the Philosophy of Psychology, Vol. I*. Cambridge, MA: Harvard University Press.
- Block, N. 1990. 'Can the Mind Change the World?' In Boolos, G (ed.) *Meaning and Method: Essays in Honor of Hilary Putnam*. Cambridge: Cambridge University 1990
- Block, N and Robert Stalnaker, 1999. 'Conceptual Analysis, Dualism and the Explanatory Gap', *Philosophical Review*, forthcoming.
- Byrne, A. 'Cosmic Hermeneutics'. 1999. *Philosophical Perspectives*, Forthcoming

- Braddon-Mitchell D. and Jackson F. 1996. *Philosophy of Mind and Cognition*. Oxford: Blackwell.
- Chalmers, D. 1996. *The Conscious Mind* New York: Oxford University Press
- Chalmers, D. 1999. 'Materialism and the Metaphysics of Modality', *Philosophy and Phenomenological Research*, forthcoming.
- Chomsky, N. 1994. 'Language and Nature' *Mind* 104: 413, 1995;
- Crane, T. and Mellor, D.H. 1994. 'There is no Question of Physicalism', *Mind* 99:185, 1990;
- Feigl, H. 1967. 'The "Mental" and the "Physical"' (Minneapolis: University of Minnesota Press. Original Publication: 1958)
- Foster, J. 1982 *The Case for Idealism*. London: Routledge
- Foster, J. 1991 *The Immaterial Self: A Defence of the Cartesian Dualist Conception of Mind*. London: Routledge.
- Harman, G. 1986 *Change in View* Cambridge, Mass: MIT Press
- Harman, G. 1990. 'The Intrinsic Quality of Experience', *Philosophical Perspectives*, 4:31-52.
- Jackson, F. 1982 'Epiphenomenal Qualia' *Philosophical Quarterly*, 32: 127-36
- Jackson, F. 1986 'What Mary Didn't Know', *Journal of Philosophy* 83 pp. 291-5
- Jackson, F. 1998 *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Clarendon
- Johnston, M. 1992 'How to Speak of the Colors', *Philosophical Studies*, 68, 1992: 221-263.
- Kim J. 1993 *Mind and Supervenience* Cambridge: Cambridge University Press.
- Kim, J. 1998 *Mind in a Physical World* Cambridge: Cambridge University Press.
- Kripke, S 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Kripke, S 1982: *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Basil Blackwell: Oxford.
- Lewis, D. 1970 'How to Define Theoretical Terms' *Journal of Philosophy* 67:427-46.
- Lewis, D. 1994 'Reduction of Mind'. In Guttenplan, S 1994 (ed) *A Companion to the Philosophy of Mind* Oxford: Blackwell.
- Loar, B. 1997. 'Phenomenal States'. In Block et al (eds.), *The Nature of Consciousness: Philosophical Debates*, Cambridge, MA: MIT Press.
- Loar, B. 1999. 'David Chalmers' *The Conscious Mind*', *Philosophy and Phenomenological Research*, forthcoming.
- Lockwood, M. 1989. *Mind, Brain and Quantum*. Oxford. Basil Blackwell.
- Lockwood, M. 1992 'The Grain Problem'. In Robinson, H. (1992) (ed) *Objections to Physicalism*. Oxford: Oxford University Press.
- Maxwell, G. 1978. 'Rigid Designators and Mind-Brain Identity'. In Savage, C (ed) 1978 *Perception and Cognition: Minnesota Studies in the Philosophy of Science Vol.9*, Minneapolis: Minnesota University Press.
- McGinn, D. 1989. 'Can we Solve the Mind-Body Problem', *Mind* 98: 349-66.

- Moore, G.E. 1922. *Philosophical Studies*. Routledge: London.
- Nagel, T. 1974. 'Armstrong on the Mind', *Philosophical Review*, 3:
- Nagel, T. 1974. 'What is it like to be a bat', *Philosophical Review*, 4: 435-50.
- Russell, B. 1927. *The Analysis of Matter* London: Kegan Paul.
- Shoemaker, S. 1994. 'Phenomenal Character', *Nous*, 28 (1994), pp.21-38
- Shoemaker, S. 1999. 'On David Chalmers' *The Conscious Mind*', *Philosophy and Phenomenological Research*, forthcoming.
- Smart J.J.C. 'The Content of Physicalism', *Philosophical Quarterly* 28 (1978) pp. 239-41.
- Smith, M. and Stoljar, D. 1998 'Global Response-Dependence and Noumenal Realism', *The Monist*, Jan. 1998, Vol.81, No.1.
- Unger, P. 1998. 'The Mystery of the Physical and the Matter of Qualities', *Midwest Studies in Philosophy*, Vol. 22.
- Yablo, S. 1992. 'Mental Causation', *The Philosophical Review*, April 1992, pp. 245-280.
- Yablo, S. 1999. 'Concepts and consciousness', *Philosophy and Phenomenological Research*, forthcoming.

Notes:

(1) It is important to note that I am merely trying to state the considerations in favor of (1) here, not defend them. Likewise the considerations in favor of (2-4) I consider in a moment.

(2) For a statement of this argument, and for an analysis of the notion of a priori entailment, see Jackson 1998, and Chalmers (1996, 1999); for criticism, see Block and Stalnaker 1999, Byrne 1999, Loar (1997, 1999), and Yablo 1999. One should perhaps talk more correctly of mental truths being a priori entailed by physical and *topic neutral* truths, but I will leave this extension largely implicit in what follows.

(3) By 'qualia', I mean the properties of experiences in virtue of which there is something it is like (in the phrase made famous by Nagel) to have those experiences.

(4) For a recent statement of this argument, see Braddon Mitchell and Jackson 1996 and Chalmers 1996. Of course, there are other reasons for resisting a priori physicalism-for example, reasons having to do with intentionality-but I will limit my discussion here to qualia.

(5) For a recent statement of this argument, see Kim 1999; for criticism, see Yablo 1992. The question of what the relation is between the two parts of causal closure-the thesis about events and the thesis about properties-is a question about the metaphysics of causation, and in particular about the metaphysics of causal relata, which I will set aside here. It is worth noting also that causal closure as I have defined it rules out the possibility of overdetermination-I will set aside this issue also.

(6) For a recent statement of this sort of argument, see Shoemaker 1999 and Chalmers 1996; for criticism, see Jackson 1982 and Chalmers 1996, 1998.

(7) It is interesting to note that Gilbert Harman's 1986 sharp distinction between what follows from what and what one ought to believe provides yet another possibility here. More particularly, in Harmanesque fashion, one might suppose that, even if (1-4) *are* contradictory, it is nevertheless a rational strategy to resolve to believe all of them so long as one also resolves not to exploit this contradiction in one's reasoning. This suggestion seems to me to present not a resolution of the puzzle posed by (1-4) so much as a way in which one might live with oneself in the absence of such a resolution, but in any event I will set it aside here.

(8) For examples of those who deny that one can have propositional knowledge of qualia, see Lewis 1994, and references therein; for strong necessities, see Yablo 1999 and Loar 1999.

(9) The theory-based conception bears some relation to the notion of physicalism discussed in Meehl and Sellars 1956 and in Feigl 1965; more explicit defense is found in Smart 1978, Lewis 1994, Braddon-Mitchell and Jackson 1996, and Chalmers 1996. There is of course the threat that a formulation of physicalism which utilizes the theory-based conception will be trivial: if the notion of a physical theory is sufficiently unconstrained, any property including irreducibly mental properties might be such that physical theory tells us about them. (For this sort of criticism, see Crane and Mellor 1990, and Chomsky 1994.) There are a number of ways in which one might seek to constrain the notion to meet this threat: by speaking of physical theory sufficiently similar to current physical theory; by speaking of physical theory sufficiently similar to commonsense physical theory; or by speaking of physical theory as constrained by the methodology of physics. I will assume here that some such strategy is available, but it will not matter for our purposes to decide which is the best.

(10) The best examples of philosophers who operate with the object-conception of the physical are Meehl and Sellars 1956 and Feigl 1965; it is also a position that one encounters regularly in discussion. There is of course the threat that a formulation of physicalism which utilizes the object-based conception will be trivial: if the notion of a paradigm physical object is sufficiently unconstrained, any property including irreducibly mental properties might be such that paradigm physical objects have them. There are a number of ways in which one might seek to constrain the notion to meet this threat: by speaking of physical objects-as-we-currently-conceive-them; by insisting that the notion of a physical object presupposes that such objects cannot turn out to be irreducibly mental; by operating with the notion of a purely physical object, where a purely physical object is something completely non-mental. As with the theory-based conception, I will assume here that some such strategy is available, but it will not matter for our purposes to decide which is best.

(11) Two other issues deserve to be mentioned briefly. First, one might object that both conceptions are inadequate because they are circular, i.e., both appeal to the notion of something physical (a theory or an object) to characterize a physical property. The response to this is that circularity is only a problem if the conceptions are interpreted as providing a reductive analysis of the notion of the physical rather than simply an understanding of it. But there is no reason why they should be interpreted in the former way. Second, it might be thought that notion of an o-physical property is open to the following objection discussed by Ned Block: "it is conceivable that there are physical laws that 'come into play' in brains of a certain size and complexity, but that nonetheless these laws are 'translatable' into physical language, and that, so translated, they are clearly physical laws (though irreducible to other physical laws). Arguably, in this situation, physicalism could be true-though not according to [this] account of 'physical property' (1980, n.4) However, at least as developed here, the object-based conception does not face this objection because the properties and laws that Block is describing *supervene* on properties required in an account of paradigmatic physical objects, and so o-physicalism *would* be true in the case he is envisaging.

(12) There is of course a large literature in support of the thesis that physical theory tells us only about dispositional properties, a literature which has at least three sources: Russell's (1927) discussion of the nature of physical theory; the approach to the structure of scientific theories and theoretical terms due to Ramsey, Carnap and Lewis (see, e.g. Lewis 1970); and an epistemological thesis that, in perception, we are acquainted only with dispositional properties of physical objects (see, e.g. Armstrong 1961, 1968). I will not in this paper be able to explore this literature or defend the thesis in any detail. My reason is partly space and partly that the contemporary exponents of the anti-physicalist position (or related positions) agree with the thesis, so there is nothing problematic about making this assumption in the course of defending physicalism; see, e.g., Chalmers 1996, pp. 153-4.

It is worth noting also that it is an oversimplification to say that physical theory tells us only about dispositional properties: physical theory also tells us about what might be called structural properties, i.e. geometrical, spatiotemporal or causal properties. But this complication does not matter for our purposes. The crucial point for our purposes is that there are categorical or non-dispositional properties of physical objects which physical theory does not tell us about.

(13) There is of course a large literature concerning the thesis that dispositions require categorical grounds, but I will not here explore this literature or defend the thesis in any detail. (For a recent defense, see Smith and Stoljar 1997.) Once again my reason is partly space and partly that contemporary defenders of the anti-physicalists position agree with the thesis; see, e.g. Chalmers 1996, n.29 on p. 375. To avoid confusion, however, it is worth noting two points: (i) The thesis does not require that if a thing has a dispositional property then there must be a non-dispositional property of *that very thing* such that the latter is metaphysically sufficient for the instantiation of the latter, i.e., it is perfectly consistent with the thesis that the categorical properties on which the dispositional properties *supervene* might be properties of

constituents of the thing in question rather than the thing itself. (ii) The thesis is a metaphysical thesis, rather than an explanatory one, i.e., it is perfectly consistent with the thesis that to explain the presence of a certain dispositional property one might cite a further dispositional property. All that is being urged is the metaphysical claim that the dispositional supervenes on the non-dispositional.

(14) For a very similar notion, see the discussion of 'Kantian Physicalism' in Jackson 1998.

(15) While I will for the most part speak of 'o-physical properties which are not t-physical', I will sometimes speak only of 'o-physical properties'. It will be clear from context whether what is intended is the class of o-physical properties *tout court*-i.e. a class which includes t-physicals-or the class of o-physicals which are not t-physicals.

(16) The other main argument against a priori physicalism is the conceivability argument. I think the strategy of this paper also holds good against this argument but I will not discuss this issue here.

(17) For a catalogue and criticisms of the other main attempts to defeat the knowledge argument, see Braddon Mitchell and Jackson 1996. They do not consider the kind of proposal that we are considering here.

(18) It is important to emphasize that the views at issue here are Russell-*inspired* rather than Russell's. The question of what Russell's actual views are is a difficult scholarly one that I will not address here. Similarly, it is important to emphasize that while there is much similarity between the position discussed in the text and that of Maxwell and Lockwood-for example, Maxwell describes his position as 'nonmaterialist physicalism' (1978; p.365) which echoes the distinction I have drawn between the theory-based and object-based conception of the physical-neither of these writers develop or defend the position in the way I will do here. For further discussion of Russell-inspired views and related matters, see Chalmers 1996, Foster (1982, 1994) and Unger 1998.

(19) It is worth emphasis that the first part of o-physicalism is neutral on the issue of whether qualia supervene on o-physical properties which *exclude* t-physical properties or on o-physical properties which *include* t-physical properties. The general question of the role of t-physical properties in o-physicalism is an interesting question, but it is not a question that I will decide here.

(20) Indeed, it is for this reason that Chalmers takes himself to have argued *against* physicalism in *The Conscious Mind*. As far as the metaphysics of mind goes, the overall conclusion of that book is disjunctive: either epiphenomenalism is true, or interactionist dualism is true, or the Russell-inspired position is true. Given Chalmers' assumption that the third disjunct is not a version of physicalism, *The Conscious Mind* is an anti-physicalist book. However, if the argument of this paper is correct, there is no reason to suppose

that the Russell-inspired position *must* be non-physicalist. It follows-contrary to what I take to be Chalmers' aim-that there is a version of physicalism which is compatible with *The Conscious Mind*.

(21)It is easy to see that an extension of this argument would show that if a proponent of causal closure wants to defend a thesis that is compatible with the idea that we discussed in §4-that if dispositional properties are causally efficacious, their categorical grounds are-he or she must likewise adopt the object-based conception of the physical, or some equivalent conception.

(22)For very different responses to the grain problem, see Maxwell 1978, Lockwood 1992 and Chalmers 1996.

(23)Material drawn from this paper was presented to the philosophy department at the University of Colorado at Boulder, and to the Consciousness and Metaphysics in Philosophy of Mind Conference, Chateau Hotel, Sydney. I am very much indebted to all who took part on those occasions. I would also like to thank the following for comments, discussion and advice: George Bealer, Helen Beebee, Andrew Botterell, David Braddon-Mitchell, Alex Byrne, David Chalmers, Cian Dorr, John Fisher, Frank Jackson, Michael Griffin, Benj Hellie, Jakob Hohwy, Stephen Leeds, Michael Martin, Daniel Nolan, Graham Oddie, Philip Pettit, Kieran Setiya, Michael Smith, Sheldon Smith, Robert Stainton, Jason Stanley, Zoltan Szabo, and Adam Vinuesa. Particular thanks to David Chalmers for many extremely enjoyable and helpful discussions about the topic of this paper.