

2002. 'The Knowledge Argument Against Dualism', *Theoria* Vol. LXIII, pp. 205-223.

The Knowledge Argument Against Dualism

by

YUJIN NAGASAWA

Australian National University

Abstract

Paul Churchland argues that Frank Jackson's Knowledge Argument against physicalism is so strong that if it defeated physicalism it would, at the same time, defeat 'substance dualism'. The purpose of this paper is to articulate this 'parity of reasons' objection. In the first part of the paper, I discuss Churchland's argument. I demonstrate that although his formulation of the objection is not wholly satisfactory, it may be revised so that the Knowledge Argument would defeat a certain form of dualism. In the second part, I apply the parity of reasons objection to David Chalmers' dualism. Chalmers rejects physicalism on the basis of the Knowledge Argument and introduces two possible forms of dualism. I show that of those two forms of dualism, Chalmers has to endorse the one that he *does not* prefer because the other is vulnerable to the parity of reasons objection.

1. Introduction

Frank Jackson's Knowledge Argument (1982, 1986), which purports to show that there can be no physicalist account of phenomenal consciousness, is one of the most famous and provocative thought experiments in the philosophy of mind. For the last twenty years, physicalists have criticised the Knowledge Argument from many different perspectives, but have reached no consensus as to exactly what, if anything, is wrong with it. Once we accept the argument, physicalism appears hopeless and there seems no choice other than dualism. The Knowledge Argument, however, might not give dualists cause to rejoice, after all. According to what I call the 'parity of reasons' objection, which is introduced by Paul Churchland, the Knowledge

Argument is so strong that if it served to defeat physicalism it would equally well serve to defeat 'substance dualism'.

The purpose of this two-part paper is to articulate the parity of reasons objection, which, in spite of its strength, has attracted little attention. In the first part, I examine Churchland's formulation of the parity of reasons objection. I suggest that while his formulation is not wholly satisfactory, it may be modified so that the Knowledge Argument would defeat a particular form of dualism to the exact extent that it would defeat physicalism. In the second part, I consider an application of the parity of reasons objection. David Chalmers, a well-known dualist and a proponent of the Knowledge Argument, introduces two possible forms of dualism and explicitly states his preference for one over the other. However, I demonstrate, by applying the parity of reasons objection, that his preferred option would be defeated by the Knowledge Argument to the exact extent that physicalism would be defeated. Therefore, I conclude, if he wishes to reject physicalism on the basis of the Knowledge Argument, he has to subscribe to the form of dualism which he *does not* prefer.

2. The Knowledge Argument Against Physicalism

Imagine Mary, a physically omniscient scientist who is confined to a black-and-white room. Although she has never been outside her room in her entire life, she has learned *everything* there is to know about the nature of the physical world from black-and-white books and lectures on a black-and-white television. Mary's complete knowledge includes everything about the physical facts and laws of physics, which will include causal and relational facts, and functional roles. This is the beginning of the Knowledge Argument.

Physicalism is the metaphysical thesis that, in the relevant sense, everything is physical, or as contemporary physicalists often put it, everything logically supervenes on the physical. Thus,

if physicalism is true Mary, who has complete knowledge about the physical, must have complete knowledge simpliciter.^{1[1]}

What will happen, Jackson continues, when the physically omniscient scientist Mary leaves her room and looks at, say, a ripe tomato for the first time? According to physicalism she should not come to know anything new because she is supposed to know everything about the physical world. It appears obvious, however, that she will discover something new upon her release; namely, ‘what it is like to see red’, a phenomenal feature of her visual experience.^{2[2]} This

^{1[1]} The Knowledge Argument—at least Jackson’s original formulation of the Knowledge Argument—is based on the assumption that if physicalism is true then *a priori* physicalism is true. *A priori* physicalism states that mental phenomena logically supervene on physical phenomena and that there is an *a priori* derivation from physical facts to mental facts. However, many philosophers reject *a priori* physicalism. For example, *a posteriori* physicalists, such as Block and Stalnaker (1999), argue that neither macrophysical nor mental phenomena logically supervene on microphysical phenomena and that there is not an *a priori* but only an *a posteriori* derivation from physical facts to mental facts. *A posteriori* physicalists reject the Knowledge Argument on the ground that Mary does not have to be able to make an *a priori* derivation from physical facts, which she knows in a black-and-white room, to mental facts, which she comes to know upon her release. Hence Mary’s surprise at finding out what it is like to see red is, they claim, perfectly consistent with (*a posteriori*) physicalism. I set aside the issue of *a posteriori* physicalism and use the notion of reduction that is based on *a priori* physicalism throughout the paper. This does not affect the force of the parity of reasons objection because, as we will see later, the targets of the parity of reasons objection are proponents of the Knowledge Argument, such as Jackson and Chalmers, who accept the assumption that if physicalism is true then *a priori* physicalism is true. (If they do not accept it, then they cannot undermine physicalism by the Knowledge Argument in the first place.) It is important to note that *a posteriori* physicalism has been introduced and elaborated as a response to antiphysicalist arguments like the Knowledge Argument. Physicalists need this sort of independent response to the Knowledge Argument because the parity of reasons objection is not applicable to all kinds of dualism.

^{2[2]} If we want to be precise we need to add many conditions to this thought experiment: Mary’s body must be painted completely black-and-white; Mary must not rub her eyes so that she does not experience phosphenes; Mary must not experience any colourful illusions or dreams, etc. In order to get rid of this complication, Robinson stipulates instead that the protagonist of the thought experiment is a congenitally deaf scientist (Robinson 1982, pp. 4-5; 1993, p. 159).

contradicts the physicalist assumption that Mary, prior to her release, has complete knowledge simpliciter. Therefore, Jackson concludes, physicalism is false.^{3[3]}

Jackson provides a ‘convenient and accurate way of displaying’ the Knowledge Argument:

- (a) Mary (before her release) knows everything physical there is to know about other people.
- (b) Mary (before her release) does not know everything there is to know about other people (because she *learns* something about them on her release).

Therefore,

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

(Jackson 1986, p. 293)

Since, if it is successful, the Knowledge Argument defeats physicalism, its proponents, such as Chalmers, John Foster and Howard Robinson, subscribe to dualism.^{4[4]} However, Churchland

^{3[3]} After sixteen years of defending the Knowledge Argument, Jackson announced in 1998 that he had changed his mind, stating that although the argument contained no obvious fallacy its conclusion, that physicalism is false, must be mistaken. In this paper, I am concerned only with Jackson’s original claim. It should be emphasised, however, that even since his ‘conversion’ Jackson still insists that if physicalism is true Mary must know what it is like to see red *before* she goes outside her room. That is, he still believes that if physicalism is true *a priori* physicalism is true. See Jackson (1995, 1998b, forthcoming).

^{4[4]} The important question that arises here is whether the mere distinction of the mental from the physical is really sufficient to establish dualism. One might argue that even if the Knowledge Argument showed the falsity of physicalism it would not immediately follow that dualism is true. The following passage by John Searle illustrates this point:

Dualists asked, “How many kinds of things and properties are there?” and counted up to two. Monists, confronting the same question, only got as far as one. But the real mistake was to start counting at all. ... It is customary to think of dualism as coming in two flavors, substance dualism and property dualism; but to these I want to add a third, which I will call “conceptual dualism.” This view consists in taking the dualistic concepts very seriously, that is, it consists in the view that in some important sense “physical” implies

contends that dualists who defend their position by appeal to the Knowledge Argument are on shaky ground, because the argument against physicalism may be directed, in an exactly parallel form, against substance dualism. Call this the ‘parity of reasons objection’.

3. The Parity of Reasons Objection

Churchland (1985a, 1985b, 1989) argues that if the Knowledge Argument were sound, it would prove far too much, contending that if, as Jackson says, the Knowledge Argument showed physicalism to be false it would equally show ‘substance dualism’ to be false. He defines substance dualism as the thesis that there exists a mental substance called ‘ectoplasm’, the ‘hidden constitution and nomic intricacies’ of which form mental phenomena, such as visual experiences (Churchland 1985a, p. 24; 1985b, p. 119). It seems that, just like physicalism, substance dualism is a perfect target for the Knowledge Argument because no belief about ectoplasm, its structure, function, composition, etc., appears to enable Mary to know, in a black-and-white room, what exactly it is like to see red. Churchland thus concludes, “Given Jackson’s

“nonmental” and “mental” implies “nonphysical.” Both traditional dualism and materialism presuppose conceptual dualism, so defined. (Searle 1992, p. 26)

Although Searle’s claim deserves serious consideration, it has no impact on the parity of reasons objection. For both proponents and opponents of the Knowledge Argument would agree with Jackson’s objection to Searle as follows:

Searle is right that there are lots of kinds of things. But if the thought is that any attempt to account for it all, or to account for it all as far as the mind is concerned ... in terms of some limited set of fundamental (or more fundamental) ingredients, is mistaken in principle, then it seems to me that we are being, in effect, invited to abandon serious metaphysics in favour of drawing up big lists. (Jackson 1998a, p. 4)

It should be noted, however, that Searle’s main thrust is consistent with the idea behind the parity of reasons objection. That is, if physicalism really were false then the mere introduction of an additional entity, such as mental substance, could not be a significant improvement.

antiphysicalist intentions, it is at least an irony that the same form of argument should incidentally serve to blow substance dualism out of the water^{5[5]} (Churchland 1989, p. 574).

Now we can illustrate Churchland's parity of reasons objection by providing the following case, a simple variation of the original Knowledge Argument.

4. The Knowledge Argument Against Dualism

Imagine Mark, a dualistically omniscient thirty-fifth century scientist who is confined to a black-and-white room. People at this time have not only the complete science of the physical, but also the complete science of the mental stuff 'X', which one of their ancestors discovered in the thirty-second century. The constitution and nomic intricacies of X ground all mental phenomena. Although Mark has never been outside his room in his entire life, he has learned *everything* there is to know about the nature of the physical world and of X from black-and-white books and lectures on a black-and-white television. Mark's complete knowledge includes everything about the physical facts and laws of physics, which will include causal and relational facts, and functional roles. Moreover, his knowledge of X provides explanations about our ordinary mental phenomena, such as thoughts and feelings in terms of their relations to X. What will happen, we may ask, when the dualistically omniscient scientist Mark leaves his room and looks at, say, a ripe tomato for the first time? Since Jackson's original Knowledge Argument is valid, the following argument is equally valid.

(a') Mark (before his release) knows everything physical and everything 'X-ish' there is to know about other people.

^{5[5]} It should be emphasised that, as a physicalist, Churchland *does not* accept the Knowledge Argument. While he provides several objections to the argument, his main complaint is that the Knowledge Argument equivocates on the notion of knowledge. According to Churchland, while in the first premise of the Knowledge Argument Jackson focuses on propositional knowledge in the second premise he focuses on nonpropositional knowledge. See Churchland (1985a, 1989).

(b') Mark (before his release) does not know everything there is to know about other people (because he *learns* something about them on his release).

Therefore,

(c') There are truths about other people (and himself) that escape the physicalist and X-ish stories.

Just as Mary's complete knowledge of the physical is silent about the phenomenal character of her visual experience, Mark's complete knowledge of the physical and X is silent about the phenomenal character of his visual experience. Therefore, if the Knowledge Argument were cogent, it would refute dualism based on X as completely as it would refute physicalism. It is now clear that the Knowledge Argument is much stronger than people tend to think; perhaps too strong. Of course, the Knowledge Argument itself cannot be rejected solely by pointing out that it might be too strong, but Mark's case shows that there is, at least at first glance, the parity of reasons problem for dualists. Dualists adopt the Knowledge Argument in order to reject physicalism despite the fact that the argument might equally well defeat certain forms of dualism. It is at least unfair for dualists to emphasise only the anti-physicalist aspect of the Knowledge Argument.

In sum: if the Knowledge Argument served as an argument against physicalism, it would equally well serve as an argument against a particular form of dualism.

5. Replies from Dualists

Jackson does not accept the parity of reasons objection to dualism. According to him, while it is possible to acquire complete knowledge based on physicalism in a black-and-white room, it is impossible to acquire complete knowledge based on dualism in a black-and-white room:

To obtain a good argument against dualism (attribute dualism; ectoplasm is a bit of fun), the premise in the knowledge argument that Mary has the full story according to physicalism before her release, has to be replaced by a premise that she has the full story

according to dualism. The former is plausible; the latter is not. Hence, there is no ‘parity of reasons’ trouble for dualists who use the knowledge argument. (Jackson 1986, p. 295)

One might think that the key issue here is a distinction between *substance dualism* and *property (attribute) dualism*. Churchland argues that ectoplasmic *substance dualism* is vulnerable to the Knowledge Argument as much as physicalism is. Jackson replies to him that ectoplasmic *substance dualism* is just a ‘bit of fun’ and that *property dualism*, a much more realistic option for him, can avoid the parity of reasons problem.^{6[6]}

Substance dualism is the metaphysical thesis that our world consists of two fundamentally distinct substances: the physical and the mental. According to this thesis, mental states are derived solely from the states of mental substances, which have only nomological connections to physical bodies. On the other hand, property dualism states that mental properties exist, while mental substances do not. According to this thesis, mental states are mere physical states with special mental properties, properties that are clearly distinct from physical properties. Hence, the essential difference between substance dualism and property dualism comes from what each takes as components of the mental world. Substance dualism regards the mental world as composed of mental substances and property dualism regards the mental world as composed of mental properties. While classifying dualism in this way is a common practice in the philosophy of mind, this distinction has, essentially, nothing to do with the parity of reasons objection, because it

^{6[6]} One might argue that Jackson’s concern here is not merely the distinction between substance dualism and property dualism. Perhaps his intention is simply to dismiss ectoplasmic substance dualism out of hand, as a viewpoint not even worthy of consideration, taking property dualism, instead, as a more plausible position. Then, he may be taken as arguing that there is no parity of reasons trouble because it is simply not plausible to suppose that anyone could have the full story according to property dualism. If this is the correct interpretation of Jackson’s passage, it runs into trouble of its own. If it is not plausible to suppose that anyone could have the full story according to property dualism, the position becomes a mere *ad hoc* hypothesis, with no substance of its own other than to shore up gaps in our knowledge.

makes no difference whether one identifies X with substances or properties. Mark's story could work perfectly well in either case.

Howard Robinson, another proponent of the Knowledge Argument, correctly realises the irrelevance of the distinction between substance dualism and property dualism. He claims that a certain kind of substance dualism is not vulnerable to the Knowledge Argument:

Jackson points out that [the parity of reasons objection] does not touch property dualism, which is all that the argument proves.^{7[7]} But neither does it touch a sensible substance dualism. 'Mental substance' is not something composed of 'ghostly atoms'—whatever that would mean—but something that is not *made of* anything at all. In so far as it has a structure, that structure would be entirely psychological—that is, would consist of the faculties, beliefs, desires, experiences, etc. There would be no autonomous sub-psychological stuff. Such a notion faces many problems, of course, but this is the Cartesian conception, not the ectoplasmic one; and against this conception the knowledge argument is irrelevant. (Robinson 1993, p. 183)

In the above passage Robinson shows that there are two kinds of substance dualism. According to the first, mental substance is composed of 'ghostly atoms'. Robinson implies that this kind of substance dualism would be defeated by the Knowledge Argument to the extent that physicalism would be defeated. According to the second kind, mental substance is not made of anything at all. Robinson claims that this kind of substance dualism would not be defeated by the Knowledge Argument. Now we can make a parallel claim about property dualism. While one kind of property dualism, according to which mental properties are composed of 'ghostly atomic properties', would be defeated by the Knowledge Argument to the extent that physicalism would be defeated, another kind of property dualism, according to which mental properties are not composed of anything at all, would not be defeated.

^{7[7]} Here Robinson refers to Jackson's passage quoted in the main text.

At this point it is clear that Churchland's simple claim that substance dualism is vulnerable to the Knowledge Argument and Jackson's simple claim that property dualism is not vulnerable to the argument, are both incomplete. I now introduce a new way of looking at dualism in order to distinguish clearly a kind of dualism that would be defeated by the Knowledge Argument from a kind of dualism that would not. This classification relies on the reducibility of our ordinary mental phenomena.

6. Reductive Dualism and Nonreductive Dualism

Reductive explanations are important for the scientific understanding of nature. For example, thermodynamics explains the temperature of a gas reductively, in terms of the mean kinetic energy of the constituent molecules.^{8[8]} The molecules that make up a gas are in constant motion and the temperature of a gas is a measure of the speed at which they move. The faster they move, the higher the temperature. Again, meteorology explains lightning reductively in terms of electric discharge. Electric discharge occurs as a result of the separation of positively and negatively charged particles in storm clouds, and lightning occurs when the power of attraction between positive and negative particles increases to a certain point. In general, the physical sciences provide reductive explanations of higher-level physical phenomena in terms of their underlying lower-level physical phenomena.

Similarly, dualists may suppose that we can reductively explain higher-level mental phenomena, like thoughts and feelings, in terms of their underlying lower-level mental phenomena, of which those ordinary mental phenomena are comprised. Call the form of dualism that adopts this kind of reductive explanation 'reductive dualism'. On the other hand, other dualists may suppose that there is no reductive explanation whatsoever for our ordinary mental

^{8[8]} It is often said that temperature (in general) is reducible to the kinetic energy of the constituent molecules of the object whose temperature is at issue. However, strictly speaking, this is mistaken. It is true only for a gas, and not for a solid. See Churchland (1996), p. 41.

phenomena. According to this form of dualism, no matter how far our sciences may advance, those mental phenomena will remain irreducible, perhaps because they are fundamental primitives of the universe. Call this kind of dualism ‘nonreductive dualism’. I now examine those two forms of dualism and show that reductive dualism would be defeated by the Knowledge Argument to the extent that physicalism would, while nonreductive dualism would not.

Reductive dualism is, in a sense, an elegant hypothesis because it presents a symmetry of the mental and physical worlds. Higher-level physical phenomena are reducible only to their underlying lower-level physical phenomena, and higher-level mental phenomena are reducible only to their underlying lower-level mental phenomena. The physical and mental worlds are clearly distinct and never overlap each other.^{9[9]}

Ectoplasmic substance dualism, which Churchland introduces, represents one kind of reductive dualism. It explains our ordinary mental phenomena, such as visual experiences, in terms of their underlying mental substance, ectoplasm, of which those mental phenomena are comprised. Just as the temperature of a gas is fully explained in terms of kinetic energy of the

^{9[9]} Since dualism is realism about *two* distinct kinds of substances or properties, the physical and the mental, one might think that there are two possible claims that reductive dualists may hold:

(R1) Our ordinary mental phenomena are reducible to their underlying *physical* phenomena, out of which those phenomena are composed.

(R2) Our ordinary mental phenomena are reducible to their underlying *mental* phenomena, out of which those phenomena are composed.

However, (R1) is not an option for dualists because it says essentially that there are no mental phenomena over and above physical phenomena and that everything is ultimately explained in terms of the physical. It is, rather, a form of *physicalism*. Therefore, all reductive dualists must hold (R2).

It is also possible for reductive dualists to claim that not only higher-level mental phenomena, but also some physical phenomena, are reducible to lower-level mental phenomena, or that not only higher-level physical phenomena but also some mental phenomena are reducible to lower-level physical phenomena. However, no dualists should want to hold this idea because it would violate fundamental conditions of acceptability on what would constitute a good explanation in both science and metaphysics; for example, simplicity, elegance and parsimony.

constituent molecules, or lightning in terms of electric discharge, our thoughts or feelings are, according to this doctrine, fully explained in terms of ectoplasm.

If we replace *X* with ectoplasm, we can see that ectoplasmic substance dualism would be defeated by the Knowledge Argument to the same extent that physicalism would. In the same manner, we can construct the Knowledge Argument against any form of reductive dualism. We replace *X* with mental substances or mental properties, out of which our ordinary mental phenomena are composed, and then let Mark learn, in addition to the knowledge afforded by the physical sciences, every reductive explanation provided by reductive dualism. Just as in Mary's case, Mark would come to know, if the Knowledge Argument were successful, something new upon his release.

It is worth emphasising again that the distinction between property dualism and substance dualism does not play a crucial role here, since it makes no difference whether higher-level mental phenomena are composed of underlying lower-level mental *substances*, like ectoplasm, or underlying lower-level mental *properties*. The Knowledge Argument against dualism above would be perfectly applicable to both reductive substance dualism and reductive property dualism to the exact extent that the original Knowledge Argument would be applicable to physicalism.

It is also worth emphasising that I am not here merely claiming that we can reject reductive dualism because the Knowledge Argument is cogent. The cogency or otherwise of the Knowledge Argument is an interesting but completely separate issue. I am, rather, making the conditional claim that *if* the Knowledge Argument successfully defeated physicalism, it would equally successfully defeat reductive dualism. For as far as the Knowledge Argument is concerned, reductive dualism is exactly parallel to physicalism. If reductive dualists rejected physicalism by the Knowledge Argument, physicalists could reject reductive dualism by the argument as well. Conversely, if reductive dualists eliminated the force of the Knowledge Argument, physicalists could eliminate its force as well. (For example, if reductive dualists are allowed to say that the Knowledge Argument against reductive dualism fails because Mark

cannot acquire complete dualistic knowledge in a black-and-white room, which dualists tend to say, then physicalists are equally allowed to say that the Knowledge Argument against physicalism fails because Mary cannot acquire complete physical knowledge in a black-and-white room. I will come back to this point in the next section.) The upshot is that, *as regards physicalism and reductive dualism, there is no reason to favour one or the other as far as the Knowledge Argument is concerned.*

7. Application of the Parity of Reasons Objection: Chalmers' Panprotopsychism

Dualists might argue that even if the parity of reasons objection is acceptable it has nothing to do with contemporary dualists anyway, since none of them subscribes to reductive dualism. According to this objection, while reductive dualism might have been popular in the modern period, it is no longer regarded as a tenable option, even among dualists. However, some serious contemporary dualists do subscribe, consciously or unconsciously, to reductive dualism. In this section, I apply the parity of reasons objection to Chalmers' panprotopsychism, which represents a contemporary version of reductive dualism.

In trying to establish a well-formed theory of consciousness, Chalmers introduces two options that dualists may take:

There are two ways this might go. Perhaps we might take [phenomenal] experience itself as a fundamental feature of the world, alongside space-time, spin, charge and the like. That is, certain phenomenal properties will have to be taken as *basic* properties. Alternatively, perhaps there is some *other* class of novel fundamental properties from which phenomenal properties are derived. ... [T]hese cannot be physical properties, but perhaps they are nonphysical properties of a new variety, on which phenomenal properties are logically supervenient. Such properties would be related to experience in the same way that basic physical properties are related to nonbasic properties such as [the] temperature [of a gas]. We could call these properties *protophenomenal* properties,

as they are not themselves phenomenal but together they can yield the phenomenal.

(Chalmers 1996, pp. 126-127)

Although Chalmers does not explicitly talk about reduction here, it is obvious that of the two hypotheses cited above, the former is nonreductive dualism and the latter, panprotopsychism, is reductive dualism.^{10[10]} The former states that phenomenal properties are not reducible to anything, because they are fundamental features of the world for which there is no further reductive explanation available in principle. That is, according to this thesis, phenomenal properties do not logically supervene on anything. The latter, panprotopsychism, states, on the other hand, that, just as the temperature of a gas is reducible to the kinetic energy of the constituent molecules, phenomenal properties are reducible to protophenomenal properties, of which those phenomenal properties are composed. Phenomenal properties, concerning which physicalism is completely silent, are explained in terms of protophenomenal properties. For, according to Chalmers, phenomenal properties *logically supervene* on protophenomenal properties (p. 126).

While Chalmers entertains those two possible dualist options, he explicitly admits his preference for panprotopsychism.^{11[11]}

Hill & McLaughlin say that I endorse epiphenomenalism, and that my anti-materialist argument implies epiphenomenalism. This is not strictly true. In fact my preferred position on the mind-body problem ... is not epiphenomenalism but the

^{10[10]} Another doctrine that might represent contemporary reductive dualism is John Eccles' interactionist dualism. According to Eccles, a given mental event is composed of millions of 'psychons', which correspond to what Descartes and Hume call an idea. Psychons interact, Eccles says, with dendrons, collections of dendrites in the cerebral cortex. See Eccles (1994), and Dennett (1991), p. 37.

^{11[11]} It seems that Chalmers was initially inclined towards nonreductive dualism but changed his mind at some point. Now he is much more sympathetic to panprotopsychism than nonreductive dualism. Compare, for example, Chalmers (1995) with his (2002). Recently, he contends that on one interpretation it is even possible to regard his panprotopsychism as a form of reductive monism.

‘panprotopsychist’ (or ‘Russellian’) position on which basic physical dispositions are grounded in basic phenomenal or protophenomenal properties. (Chalmers 1999, p. 492)

In the following, I demonstrate that the parity of reasons objection presents Chalmers with a dilemma. As long as he wants to use the Knowledge Argument to undermine physicalism he has to give up panprotopsychism, which he prefers, and endorse nonreductive dualism, which he does not prefer. Conversely, if he wishes to endorse panprotopsychism, he must relinquish his appeal to the Knowledge Argument.

Chalmers rejects physicalism by appeal to the Knowledge Argument, which he thinks successfully demonstrates the ‘failure of logical supervenience’ (Chalmers 1996, p. 140). However, panprotopsychism, to which he adheres, is an exact parallel of physicalism as far as the Knowledge Argument is concerned. While physicalism says that phenomenal properties logically supervene on physical properties, panprotopsychism says that phenomenal properties logically supervene on protophenomenal properties (p. 126). Because of their parallel structure, the Knowledge Argument defeats panprotopsychism to the exact extent that it defeats physicalism. We replace X in Mark’s case with protophenomenal properties, and then let Mark learn, in addition to the knowledge afforded by the physical sciences, every reductive explanation provided by the complete science of protophenomenal properties. Again, Mark would come to know, if the Knowledge Argument were successful, something new upon his release. We could simply reject Chalmers’ reductive dualism as a result of this consequence if the Knowledge Argument really defeated physicalism.

Now I consider two possible objections that Chalmers might raise against my argument.

Objection 1: Panprotopsychism is Not Reductionist

Chalmers might claim that his panprotopsychism is not reductionist, on the grounds that even if phenomenal properties are reducible to protophenomenal properties, protophenomenal properties themselves are not reducible to anything. Hence, he might conclude, Mark cannot have *complete*

knowledge of protophenomenal properties and, contrary to physicalism, panprotopsychism would not be undermined by the Knowledge Argument.

This reply is vulnerable to another parity of reasons objection. If dualists are allowed to reject the Knowledge Argument simply by saying that Mark does not have *complete* dualistic knowledge, physicalists are equally allowed to reject the Knowledge Argument simply by saying that Mary does not have *complete* physical knowledge. For, as Chalmers himself admits, certainly there are fundamental physical primitives such as space-time, spin or charge, that are not explained reductively by the physical sciences. It is clear that Mary cannot reductively explain what space-time, spin or charge are. Thus, to the extent that Chalmers could escape the consequences of the Knowledge Argument, physicalists could escape its consequences too.

Obviously, this reply expects too much of reductive explanations. Although there are many basic irreducible properties in the physical world, the physical sciences have successfully explained higher-level physical phenomena in terms of their underlying lower-level physical phenomena. It is hard to see why Chalmers' panprotopsychism does not work similarly if phenomenal properties do logically supervene on protophenomenal properties and if protophenomenal properties are related to phenomenal properties "in the same way that basic physical properties are related to nonbasic properties such as [the] temperature [of a gas]" (pp. 126-127). Changing the topic from phenomenal properties to protophenomenal properties does not save Chalmers' dualism.

Objection 2: Mark Cannot Learn About Protophenomenal Properties in a Black-and-White Room

Chalmers might also argue that panprotopsychism is irrelevant to the Knowledge Argument because although phenomenal properties are reducible to protophenomenal properties, Mark cannot know what protophenomenal properties are in a restricted environment. That is, protophenomenal properties of colour experiences are not something learnable in a black-and-

white room.^{12[12]} If this response were right, we would not be able to apply Mark's case to panprotopsychism. However, this objection is not compelling.

Chalmers contends that protophenomenal properties “are not themselves phenomenal” (Chalmers 1996, p. 127). (Imagine how absurd it would be to say, for example, “I had a protophenomenal experience of a red sensation yesterday”!). Protophenomenal properties are supposed to be, rather, fundamental constituents of phenomenal properties that are necessary for reductive explanations of phenomenal properties. Chalmers also says that combining protophenomenal properties yields phenomenal properties (p. 127), which implies that each phenomenal property is identified by the composition of its underlying protophenomenal properties. From these characteristics of protophenomenal properties there seems no reason to suppose that panprotopsychism requires Mark actually to see a red object in order to know what it is like to see red. If panprotopsychism is true then, just as Mary acquires complete physical knowledge, Mark should be able to acquire complete dualistic knowledge, and come to know what it is like to see red in a black-and-white environment.

Of course, Chalmers might choose to accept all this and still *stipulate* that Mark has to have a relevant experience in order to understand its underlying protophenomenal properties. However, it is hard to understand the motivation for doing that when, again, protophenomenal properties “are not themselves phenomenal” (p. 127). It is unclear how having a particular experience helps in understanding protophenomenal, that is *nonphenomenal*, properties. Chalmers says that “it is very hard to imagine what a protophenomenal property could be like” (p. 127), but he cannot just stipulate characteristics of protophenomenal properties without providing good reasons.

Chalmers argues that the physical sciences cannot solve the ‘hard problem’ of phenomenal consciousness by saying that “[p]hysical explanation is well suited to the explanation of *structure* and of *function* ... [b]ut the explanation of consciousness is not just a matter of explaining

^{12[12]} Notice that this response resembles Jackson's reply to Churchland's parity of reasons objection.

structure and function” (p. 107). However, as we have seen, assuming that panprotopsychism is intelligible at all, it can merely provide structural and functional explanations of phenomenal properties in terms of their underlying protophenomenal properties.

Notice, ironically, that Chalmers’ panprotopsychism may here be parallel to physicalism, thus leading again to trouble in the form of parity of reasons. Chalmers argues that protophenomenal properties are not themselves phenomenal, but that together they yield the phenomenal (p. 127). However, physicalists may equally argue that the physical constituents of the brain are not themselves phenomenal, but that together they yield the phenomenal.

Chalmers is a well-known proponent of the Knowledge Argument^{13[13]}, but if he endorses panprotopsychism he is not entitled to use the argument in the way he does. In other words, if he wants to reject physicalism on the basis of the Knowledge Argument, he has to endorse nonreductive dualism, which states that “conscious experience [is] a fundamental feature, irreducible to anything more basic” (Chalmers 1995, p. 337). However, the problem is that Chalmers thinks, as we have seen, nonreductive dualism is less plausible than panprotopsychism. At this point, therefore, the parity of reasons objection presents Chalmers with a dilemma: He has to either (1) hold onto panprotopsychism and give up the Knowledge Argument or (2) hold onto the Knowledge Argument and give up panprotopsychism. Obviously, Chalmers would not want to do either of these.

8. Conclusion

The Knowledge Argument has been welcomed by dualists as one of the strongest motivations for rejecting physicalism and endorsing dualism. However, as we have seen, the parity of reasons objection shows that, as far as the Knowledge Argument is concerned, reductive dualism is no

^{13[13]} For Chalmers’ defence of the Knowledge Argument, see Chalmers (1996), pp. 140-146.

more advantageous than physicalism. Why then, has the Knowledge Argument been so vigorously supported by dualists and opposed by physicalists?

Jackson contends that the Knowledge Argument is based on an antiphysicalist intuition that “there are certain features of bodily sensations ... which no amount of purely physical information includes” (Jackson 1982, p. 127). However, this intuition seems to be based on a more basic intuition: there are certain features of bodily sensations which no amount of intelligible, reductive explaining can include.^{14[14]} Clearly, by itself, this is not an intuition about physicalism. Perhaps the reason it has been taken for granted by both dualists *and* physicalists that the Knowledge Argument is an argument against physicalism is the following: physicalists are the ones who have most ambitiously and eagerly tried to provide intelligible, reductive explanations of phenomenal consciousness in the last couple of decades. However, it seems to me, providing this sort of explanation is not only a necessary condition for the completion of the physicalist project, but also for that of any alternatives.^{15[15]}

References

Alter, T. (1998): ‘A Limited Defence of the Knowledge Argument’, *Philosophical Studies* 90, pp. 35-56.

^{14[14]} For similar claims see Churchland (1989), pp. 573-574, and Alter (1998), pp. 49-51.

^{15[15]} Versions of this paper were read at the Australian National University in 2001 and at Toward a Science of Consciousness in Tucson in 2002. I would like to thank everyone in the audiences. I would also like to thank the Philosophy Program at the Australian National University for the financial support. I am grateful to Torin Alter, Stephen Biggs, Dave Chalmers, Daniel Cohen, Nic Damnjanovic, Frank Jackson, Josh Parsons, Howard Robinson, Laura Schroeter, and Kim Sterelny for helpful discussion. I am greatly indebted to Martin Davies and two anonymous referees for *Theoria* for substantial help in a number of critical places. Finally, I am very grateful to Daniel Stoljar for insightful, generous comments and encouragement

- Alter, T. (1999): 'The Knowledge Argument', *A Field Guide to the Philosophy of Mind*, (<http://host.uniroma3.it/progetti/kant/field/ka.html>).
- Alter, T. (2001): 'Know-How, Ability, and the Ability Hypothesis', *Theoria* 67, pp. 229-239.
- Bigelow, J. and R. Paragetter (1990): 'Acquaintance with Qualia', *Theoria* 61, pp. 129-147.
Reprinted in P. Ludlow, Y. Nagasawa and D. Stoljar (forthcoming).
- Blackburn, S. (1992): 'Filling in Space', *Analysis* 50, pp. 62-65.
- Block, N., O. Flanagan, and G. Güzeldere (eds.) (1997): *The Nature of Consciousness*, Cambridge, MA: MIT Press.
- Block, N. and R. Stalnaker (1999): 'Conceptual Analysis, Dualism, and the Explanatory Gap', *Philosophical Review* 108, pp. 1-46.
- Chalmers, D. J. (1995): 'The Puzzle of Conscious Experience', *Scientific American* 237, pp. 80-86. Reprinted in P. van Inwagen and D. W. Zimmerman (eds.) (1998): *Metaphysics: The Big Questions*, Oxford: Blackwell, pp. 333-341. Page references to reprinting.
- Chalmers, D. J. (1996): *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford University Press.
- Chalmers, D. J. (1999): 'Materialism and the Metaphysics of Modality', *Philosophy and Phenomenological Studies* 59, pp. 473-496.
- Chalmers, D. J. (2002): 'Consciousness and Its Place in Nature' in S. Stich and T. Warfield (eds.) (2002): *Blackwell Guide to Philosophy of Mind*. Oxford: Blackwell.
- Churchland, P. M. (1985a): 'Reduction, Qualia, and the Direct Introspection of Brain States', *Journal of Philosophy* 82, pp. 8-28.
- Churchland, P. M. (1985b): 'Review of Robinson's *Matter and Sense*', *Philosophical Review* 94, pp. 117-120.
- Churchland, P. M. (1989): 'Knowing Qualia: A Reply to Jackson', in his (1989): *A Neurocomputational Perspective*, Cambridge, MA: MIT Press, pp. 67-76. Reprinted in N. Block, O. Flanagan, and G. Güzeldere (1997), pp. 571-577. Page references to reprinting.

- Churchland, P. M. (1996): *Matter and Consciousness*, Cambridge, MA: MIT Press.
- Dennett, D. C. (1991): *Consciousness Explained*, Boston: Little Brown and Company.
- Eccles, J. C. (1985): 'Mental Summation: The Timing of Voluntary Intentions by Cortical Activity', *Behavioral and Brain Sciences* 8, pp. 542-543.
- Eccles, J. C. (1994): *How the Self Controls Its Brain*, Berlin: Springer-Verlag.
- Endicott, R. P. (1995): 'Refutation by Analogous Ectoqualia', *Southern Journal of Philosophy* 33, pp. 19-30.
- Foster, J. (1991): *Immaterial Self*, London: Routledge.
- Jackson, F. (1982): 'Epiphenomenal Qualia', *Philosophical Quarterly* 32, pp. 127-136. Reprinted in P. Ludlow, Y. Nagasawa and D. Stoljar (forthcoming).
- Jackson, F. (1986): 'What Mary Didn't Know', *Journal of Philosophy* 83, pp. 291-295. Reprinted in P. Ludlow, Y. Nagasawa and D. Stoljar (forthcoming).
- Jackson, F. (1994): 'Finding the Mind in the Natural World', in R. Casati, B. Smith and G. White (eds.) (1994): *Philosophy and the Cognitive Sciences: Proceeding of the 16th International Wittgenstein Symposium*, Vienna: Verlag Holder-Pichler Tempsky, pp. 101-112. Reprinted in N. Block, O. Flanagan, and G. Güzeldere (1997), pp. 483-492.
- Jackson, F. (1995): 'Postscript' in P. K. Moser and J. D. Trout (eds.) (1995): *Contemporary Materialism*, London: Routledge, pp. 184-189. Reprinted in P. Ludlow, Y. Nagasawa and D. Stoljar (forthcoming).
- Jackson, F. (1998a): *From Metaphysics to Ethics*, Oxford: Oxford University Press.
- Jackson, F. (1998b): 'Postscript on Qualia', in his (1998): *Mind, Method and Conditionals*, London: Routledge, pp. 76-79. Reprinted in P. Ludlow, Y. Nagasawa and D. Stoljar (forthcoming).
- Jackson, F. (forthcoming): 'Looking back on the Knowledge Argument', in P. Ludlow, Y. Nagasawa and D. Stoljar (forthcoming).

- Jacquett, D. (1995): 'The Blue Banana Trick: Dennett on Jackson's Color Scientist', *Theoria* 61, pp. 217-230.
- Ludlow, P., Y. Nagasawa and D. Stoljar (eds.) (forthcoming): *There's Something About Mary: Essays on Frank Jackson's Knowledge Argument*, Cambridge, MA: MIT Press.
- Robinson, H. (1982): *Matter and Sense*, Cambridge: Cambridge University Press.
- Robinson, H. (1993): 'The Anti-Materialist Strategy and the "Knowledge Argument"', in his (ed.) (1993): *Objections to Physicalism*, Oxford: Oxford University Press, pp. 159-183.
- Searle, J. (1992): *The Rediscovery of the Mind*, Cambridge, MA: MIT Press.
- Stoljar, D. (2000): 'An Argument for Neutral Monism', Manuscript.
- Stoljar, D. (2001a): 'Physicalism', *Stanford Encyclopedia of Philosophy* (<http://plato.stanford.edu/entries/physicalism/>).
- Stoljar, D. (2001b): 'Two Conceptions of the Physical', *Philosophy and Phenomenological Research* 62, pp. 253-281.
- Strawson, G. (1994): *Mental Reality*, Cambridge, MA: MIT Press.
- Tye, M. (1995): *Ten Problems of Consciousness*, Cambridge, MA: MIT Press.
- Tye, M. (2000): *Consciousness, Color, and Content*, Cambridge, MA: MIT Press.
-