The Ontology of Subjective Physicalism

Robert J. Howell Southern Methodist University

Two intuitions generate the core of the mind-body problem. On the one hand there is the idea that conscious mentality, with all its Technicolor vivacity, is bound to have aspects that elude the cold, objective descriptions of science. On the other hand, there is the deep sense that everything is composed of the same basic physical stuff, and that to think otherwise is to grant conscious beings like ourselves a convenient pride of place in the world's ontological catalogue. Upon reflection, however, many people are inclined to give up one or the other of these intuitions because they seem to conflict. On some level one of the intuitions maintains that there is something "special" about conscious mentality, while the other maintains that it is not "special," but is crafted of the same clay as everything else.

I maintain that we can hold onto both of our basic intuitions because they do not really conflict. In particular, they can be accommodated by a view I shall call *subjective physicalism*. According to subjective physicalism, the world is completely physical, but no objective theory—including physics—can completely describe the world. In particular, there are some physical states that are subjective, in that those states must be undergone in order to be fully grasped. Despite the fact that this view embraces our most fundamental feelings about the nature of the world, there is no denying the air of paradox that it generates. I maintain, however, that this disappears as long as one maintains a rigorous separation between epistemological and metaphysical issues. Policing this border is easier said than done, however, as there are numerous points where epistemological concerns come garbed in metaphysical clothing. Even if one ultimately rejects subjective physicalism, therefore, it

highlights the way metaphysical commitments "upstream" of the mind-body problem severely impact one's stance on that issue.

Subjective physicalism can be developed in two distinct ways, involving either of the following two claims:

- 1. A full physical description of the world leaves nothing out. All properties can receive objective, physical descriptions. Nonetheless, there are some properties that cannot be grasped fully unless they are grasped subjectively, via conscious experiences, as well as by objective physical descriptions.
- 2. Some physical properties can be grasped only subjectively. The properties that underwrite conscious experiences (e.g. qualia) are physical, but they are not identical with any property mentioned in a completed physics.

Call a view that accepts 1 *inclusive subjective physicalism*, and a view that accepts 2 *exclusive subjective physicalism*. According to inclusive subjective physicalism, a complete physics will refer to every property and event that there is. There are simply ways of understanding those properties that will not be imparted by an understanding of the theoretical descriptions of physics. According to exclusive subjective physicalism, on the other hand, some properties simply cannot be accessed by theoretical objective descriptions. Both exclusive and inclusive subjective physicalism have their virtues, but in this paper I will pursue only the inclusive version which accepts 1.¹ I will subsequently refer to this view simply as subjective physicalism.

Subjective physicalism is bound to be confused with other, more standard approaches to the problem of consciousness. After all, many views hold that physicalism is metaphysically correct even if there is some epistemic advantage to be gained by entering certain states.² Ironically, however, I agree with most of the criticisms of the existing views.³

In my mind, physicalists have not acknowledged the importance nor the difficulty of accomodating subjectivity within a physicalist metaphysics. The standard physicalist line essentially holds that there are two ways of knowing some physical facts, one physicalistic and scientific, and the other phenomenal. This is correct, but more has to be said to avoid dualism. In particular, it seems that given a certain metaphysics of properties and possible worlds, the "phenomenal" way of knowing states either presupposes non-physical properties or fails to accommodate what we find crucial to the subjective point of view. No number of conceptual curlicues will help the physicalist here, in part because it doesn't seem as though concepts can explain the subjective nature of experience, which seems to predate phenomenal concepts. That there is something it is like to experience conscious states is more fundamental than the fact that there is something peculiar about our knowledge of those states. Subjective physicalism acknowledges this while urging that new, non-physical properties need only be introduced if other, optional elements of the dualist's metaphysics are presupposed. Instead of introducing new concepts to explain the dualist's intuition, therefore, subjective physicalism accepts that intuition at face value but questions the metaphysical presuppositions that underwrite the inference to property dualism.

I will present the view in several stages. First, I will explain the operative notions of subjective and objective, indicating why one should believe that no objective description of the world can be complete. Next, I will propose a plausible understanding of physicalism which detaches that notion from the descriptive potential of physical theory. After a brief presentation of the position of subjective physicalism in light of these preliminary steps, I will consider two arguments that push subjective physicalism towards property dualism. This will bring deeper ontological matters to the fore, highlighting the particular notions of properties and possible worlds involved in subjective physicalism. Once these commitments

are on the table, a more systematic description of subjective physicalism will be presented, and a case will be made for its being the most appealing position available on the relationship between conscious states and physical states.

The Subjective vs. The Objective: The Real Lesson of the Knowledge Argument

Like Nagel's "What is it Like to be a Bat?", Jackson's knowledge argument captures the intuition that there is something about conscious experience that outstrips anything that science could possibly teach us about it.⁴ Where Nagel's argument leaves one unsure of its implications for physicalism, however, Jackson's provides the needed metaphysical bite. Nevertheless, I think Nagel had the right target all along—instead of arguing that physicalism is false, Jackson would have done better arguing that the world cannot be fully described by objective theories. That is, in essence, the conclusion of what I call the knowledge argument against objectivism.⁵

This slightly modified version of Jackson's argument goes as follows. Mary is a brilliant scientist who has lived her life in a black and white room. During her prolonged imprisonment she was taught all of physics, neuroscience, and biology through black and white computer screens. In fact, she eventually gained all the information about the world that could possibly be conveyed to her through such screens and monitors. At that point she had all the objective information about the world. Nevertheless, when she left the room to be presented with a red rose by her captor, she saw the red of the rose and learned something new—she learned what it is like to see red. Thus, not all information is objective information.

The most obvious modification of Jackson's argument is that it is now an argument against the claim that all information is *objective*.⁶ This is a conclusion that is not, on the face

of it, ontological. This version of the argument avoids the temptation to think that something non-physical is needed to "take up the slack" left by physical explanations of the world.⁷ That view is subject to the following *tu quoque* argument from the physicalist.

...if Jackson's argument were sound, it would prove far too much. Suppose Jackson were arguing not against materialism, but against dualism: against the view that there exists a nonmaterial substance—call it "ectoplasm"—whose hidden constitution and nomic intricacies ground all mental phenomena. Let our cloistered Mary be an "ectoplasmologist" this time, and let her know¹ everything there is to know about the ectoplasmic processes underlying vision. There would still be something she did not know²: what it is like to see red. Dualism is therefore inadequate to account for all mental phenomena!⁸

This *Just More Stuff* objection claims that the knowledge argument doesn't really lead to a separation of the mental stuff from the physical stuff, since a completed science of mental stuff wouldn't help Mary either. The anti-objectivist version of the argument insists, however, that the issue is not the type of stuff, but rather the way it is known. The reason Mary's list of the world's constituents is inadequate is not because it misses some stuff or property which could be added to the list. Mary's list is inadequate because it is a list—there are aspects of the world which are poorly served by objective depiction and can only be completely grasped by occupying a subjective state. In other words, the knowledge argument shows that the problem with physical theories is not their subject matter, but their approach to it. The problem is that they are objective, and if dualism is presented as objective in the same sense, then it is as vulnerable to the knowledge argument as physicalism.

The sense of "objective" and "subjective" presupposed by this version of the knowledge argument, as well as by subjective physicalism, can be captured by the following necessary condition for theory objectivity:

Necessary condition for theory objectivity: An objective theory cannot require that one enter or be able to enter any token state of determinate type T in order to fully understand states of type T.⁹

In the case at hand, an objective theory of a particular type of experience cannot require that one have a token of that type of experience in order to have a complete understanding of that type of experience. If we operate on the assumption that Mary can learn about the world outside of her room only by objective theories while locked in her room, and that while there she learns all any true objective theory can convey, then the fact that she still fails to understand something about the world shows that there are some states that must be entered in order to be fully understood. This does not show that they are not physical, it just shows that a complete grasp of them cannot be gained solely by objective theories.¹⁰

A full defense of this version of the knowledge argument would require a paper to itself.¹¹ The basic idea is this, however: all of the physicalist responses to Jackson's argument that grant that Mary has an "aha"-moment upon leaving the room must maintain that her epistemic achievement is a result of her becoming "hooked-up" to the world of colors in a way that she had previously only read about. By itself, however, being "hooked-up" to a process one had previously only read about is not sufficient for an epistemic gain—I could have read about the effect a salt pill has on my blood, but learn nothing by actually taking it. The only plausible "hooked-up" responses, therefore, must fill in the details with particular stories about how being hooked-up generates knowledge. But these views must, in the end, entail that the objectivity constraint is violated—they must require that there is some epistemic gain that Mary can make only in virtue of undergoing the state that she now knows about. There is, therefore, something that objective descriptions leave out.

Again, although physics is somehow incomplete, we are not yet forced to conclude that this incompleteness is ontological. So far, this version of the knowledge argument only makes a point about understanding and the descriptive potential of theories, where "descriptive potential" is determined in part by the theories' ability to lead theorists to a significantly new cognitive state. Nevertheless, there are several paths leading from this point to ontological conclusions. The first step in blocking these paths involves producing a definition of physicalism that is ontological and untainted by epistemic elements.

Defining Physicalism

The intelligibility of subjective physicalism depends upon keeping the epistemic and the metaphysical at arm's length from one another, so if physicalism is a metaphysical thesis, it must be free of epistemic elements. A thesis about the incompleteness of objective representations of the world should not automatically entail that the furniture of the world includes something other than the physical. Nevertheless, some definitions of physicalism might have that result.¹²

To avoid epistemicizing the physical, I propose a supervenience definition of physicalism: physicalism is true iff everything metaphysically supervenes upon the physical. Supervenience definitions capture the sense in which everything is completely metaphysically grounded in the physical, which is what is required by the basic monistic thrust of physicalism.¹³ Not just any supervenience thesis will do, however. Here, I will limit myself to providing my preferred definition.

SVP: Any metaphysically possible world that is a physical duplicate of our world is either a duplicate of our world simpliciter or it contains a duplicate of our world as a proper part.¹⁴

SVP is meant to capture the sense in which physicalism is a contingent thesis. Intuitively, physicalism is a claim about our world that is not falsified in virtue of strange goings on in other worlds—if there are other worlds like ours, but with ghosts on the loose, these should not falsify a physicalist thesis about our world.¹⁵ On the other hand, we cannot completely ignore worlds with furniture other than ours.¹⁶ Doing so ignores alien entities or properties that could problematize actual world connections in ways that physicalism should disallow.¹⁷ (Physicalism would intuitively be false, for example, if beliefs were necessitated by brain states only in worlds where there were no ghosts—physicalism should demand a closer relation than that.) This definition avoids both of these problems.

Implicit in SVP is a distinction between a broad sense of "physical" and a narrow sense. The former, which is what is being defined by the supervenience thesis, applies to anything that is physicalistically respectable. The latter, which appears within the thesis itself, applies to a narrower group of properties upon which all the others supervene. For the purposes of this debate I propose a negative definition of "physical" in its narrow sense. We are inclined to reject many definitions of the physical because they are apt to include paradigmatic examples of the non-physical in their extension. Two features in particular should not be basic and ineliminable in the narrow physical: phenomenality and intentionality. If a physical thing has a phenomenal property (there is something that it is like to have that property) or an intentional property (a property in virtue of which the thing represents something else) that property had better obtain in virtue of some property or properties that are not intentional or phenomenal. For this reason, I propose a negative definition of the narrow that is similar to that offered by Crook and Gillet (2001):

NIP: Something is physical iff it is fundamental, contingent, and is not phenomenal or intentional.¹⁹

The resulting notion of physicalism is strictly metaphysical. Roughly speaking, physicalism is true iff everything is metaphysically grounded in the fundamental features of the world that are themselves non-mental.

Subjective Physicalism

The basic thesis of subjective physicalism can now be made more precise. The world is completely metaphysically grounded in the physical, in that all things, properties and states supervene upon contingent things, properties and states which are not fundamentally intentional or phenomenal. Nevertheless, some of those supervenient states and properties are "subjective" in the sense that they cannot be fully grasped except by an agent that is undergoing them. Thus there is a sense in which physicalism is true, despite the fact that physics—or any other objective science, for that matter—cannot provide a complete understanding of the world.²⁰

It is very tempting to view subjective physicalism as a form of dualism. The worry is that if physics is ontologically complete there is little substantive sense to be given to Mary's learning anything upon exiting her room. Her pre-release ignorance, however, seems significant. It is an ignorance of something about the world and the minds of its denizens—an ignorance that we should sorely regret if we were in her shoes. Subjective physicalism seems to downgrade Mary's epistemic achievement to the point that has little or no significance. The dualist thus poses the following dilemma for subjective physicalism:

The Dualist's Dilemma: Either Mary comes to grasp new properties when she leaves her room, in which case property dualism is true, or she doesn't, in which case she grasps everything there is to grasp using only physical descriptions.

In fact the subjective physicalist should not be persuaded by the dualist's dilemma, because the dilemma itself makes metaphysical assumptions that the subjective physicalist is committed to denying. For one thing, even if Mary does come to grasp new properties, it is an open question whether those properties are physical or not. Exclusive subjective physicalism, for example, would maintain that they are. What is most important for the purposes of this paper, however, is that the dualist's dilemma cannot simply be dismissed by claiming that it is guilty of some sort of epistemic fallacy. It can only be overcome by questioning its underlying presuppositions about the individuation of properties and possibilities. That is, the physicalist must take a stand on metaphysical issues that are often left out of these debates. In particular, the subjective physicalist should adopt a fully extensionalist metaphysics that repudiates methods of individuating properties and possible worlds in terms of concepts or cognitive capacities.

The Presentation Argument and the Metaphysics of Properties

Any view admitting that Mary actually learns something when she leaves her room runs the risk of slipping down the slope to dualism. If subjective physicalism is to resist this slide without trivializing the knowledge Mary gains, it had better have a clear response to the arguments that have traditionally pushed philosophers down that ontological slope. The first argument, which I call The Presentation Argument, is often attributed to Max Black, but has in recent years been revived by Stephen White.

a. The Presentation Argument for Property Dualism

The simplest argument that pushes subjective physicalism towards dualism is articulated in J.J.C. Smart's early defense of the identity theory. After disposing of two less serious challenges to the identity theory, Smart considers the following challenge:

it may be possible to get out of asserting the existence of irreducibly psychic processes, but not out of asserting the existence of irreducibly psychic properties. For suppose we identify the Morning Star with the Evening Star. Then there must be some properties which logically imply that of being the Morning Star, and quite distinct properties which entail that of being the Evening Star. Again, there must be some properties (for example, that of being a yellow flash) which are logically distinct from those in the physicalist story.

Indeed, it might be thought that the objection succeeds at one jump. For consider the property of "being a yellow flash." It might seem that this property lies inevitably outside the physicalist framework…²¹

Smart's objector does a good job of expressing just how unavoidable property dualism can seem. The winnowing of one's ontological commitments by empirical discovery usually involves the recognition that what one previously thought to be two things is in fact one. When it comes to things, this strategy is effective and uncontroversial: one's mistaken impression that there were two things can be explained by the fact that one came to know about a single thing by two distinct properties yet failed to realize that they were both properties of that thing. This ontological pruning does not seem to work in the case of properties, however, for the simple reason that the explanation of the appearance of multiple properties will have to be explained by the existence of multiple properties (properties of properties, perhaps) that are responsible for those appearances. Ironically, therefore, when one makes an informative identity between properties one's ontological commitments actually increase!

This version of *the presentation problem*, does not stem from anything particular to conscious states; it is a general problem for property reduction. In the case of mental properties, one initially thinks there are two properties, one well-groomed physical property,

and one "touchy-feely" mental property known only "from the inside." After a little empirical work, one concludes that there really is only one property—the well-groomed neuroscientific property—previously known by two of its properties: its property of being a certain neuro-scientific state, and the property of constituting a certain feeling for the subject that instantiates the property. It seems one has gotten nowhere when it comes to decreasing the number of properties to which one is committed or when it comes to eliminating touchy-feely properties. At best the touchy-feely property is now a second-order property, but it seems no less troublesome for all that.

When presented in this way, the property dualist argument seems almost inescapable. It is telling, however, that the famous idiosyncrasies of qualitative consciousness are not playing much of a role here. The problem doesn't stem from the peculiarity of consciousness *per se*, but from the fact that ontological "pruning" seems to be unachievable on the property level, at least given the model of reduction that we apply to objects. In fact, the presentation problem makes a dubious assumption about the individuation of properties: it assumes that when there appear to be two properties, there must actually be two properties. Once one drops this assumption, the path is paved for ontological reduction and the dualistic argument is blocked.²³

b. The Presentation Argument and Property Intensionalism

It is crucial that we be clear on whether or not properties are individuated intensionally or extensionally. According to the intensional view, properties are individuated in part by the way we can think about them. Extreme intensionalism about properties would hold that there is a property answering to every coherent concept, and if two concepts are distinct, so then are the properties they express. Extreme intensionalism

can be regimented by specifying individuation conditions for concepts, or by appealing to possible concepts. For example, one might not want to say that the concept of being an unmarried male is distinct from that of being a bachelor. So, one might say that concepts are distinct iff *a priori* reflection could not establish that they necessarily have the same extension.²⁴ Connecting this condition to property intensionalism, one gets:

(PI) F and G are distinct properties iff *a priori* reflection alone cannot show that the concept of being F and the concept of being G are necessarily coextensive.²⁵

Intensionalism is attractive, in part because it makes the epistemology of properties straightforward. We can know about property identities and differences because they are metaphysically individuated by conceptual abilities. The extensionalist, on the other hand, insists upon mind-independent individuation conditions for properties. One appealing view is to individuate properties by the causal powers they bestow on their bearers.²⁶ While attractive, this is only one possible extensionalist view. The general position is that properties are not individuated in part by the ways that we can think about them.

The presentation argument depends upon an intensionalist conception of properties. Recall the general problem: one cannot achieve ontological parsimony by property identification, because explaining the *appearance* of multiple properties itself requires that there *be* multiple properties. The extensionalist denies the necessity of this connection. It should, of course, be granted that a fully rational individual thinks there are two properties when there is really one only when that individual has two perspectives on that property that he cannot reason between *a priori*. But it is only by using PI that we get from this to the conclusion that there must be two properties in virtue of which the

object is known. Otherwise, there being two perspectives on one property entails nothing at all about how many properties there are, second-order or otherwise.

c. Tensions within Property Intensionalism

Simply by examining the general structure of the presentation argument, we find reason to believe that intensionalism is bound to generate a profligate ontology. Recall the puzzle: when reducing properties in an attempt to prune one's ontology, intensionalism actually increases the number of properties one must recognize. Why? Because there are entailments from the appearance of property distinctness to actual property distinctness. Unless we have *a priori* assurances that there is such an entailment, we should have serious doubts about intensionalism's ability to ontologically economize.

Ontological profligacy is not the only reason to be suspicious of intentionalism, however. Intensionalism makes knowing property identities a rather simple matter because it inserts the epistemic into something that is properly metaphysical. Concepts are individuated psychologically and have to do with the way that we think about things. It is therefore counterintuitive that properties, features of the world, should be hostage to concepts. To the extent that we wish to remain realists, maintaining that the world is not of our making and has the features it does independently of our minds, we should keep concepts and properties clearly distinct and we should avoid tying them too closely to one other. This is not to say that we do not often have epistemic license to infer facts about properties based upon *a priori* conceptual reflection, but this should not be confused with a metaphysical principle of individuation.²⁷ Accepting property intensionalism is potentially a step towards anti-realism, and if it is optional, we should avoid it.²⁸

It might be complained that I am saddling the intensionalist with a more outrageous program than he needs. At least when leveling the charge of anti-realism, I have more or less assumed that the existence of the properties in question depends somehow upon the existence of the corresponding concepts. Such a view is not forced upon the defender of PI, however. Instead, it could be maintained that the concepts in question pick out independently existing properties that are not metaphysically individuated by those concepts. If so, my charge of anti-realism would be ill-founded.²⁹

Two important challenges face this realist version of PI. First, for it to be plausible, the properties that are picked out must have individuation conditions that are non-epistemic—i.e. they must have extensional individuation conditions. It is doubtful, however, that purely extensional individuation conditions can be provided, at least for the properties that make trouble in the presentation problem. These are appearance properties, and it seems that even by the intensionalist's own lights, they will have to be individuated epistemically, based on the way they appear.³⁰ Thus, it seems implausible to say that these properties are simply "picked out" and not shaped, in part, by our ways of picking them out. Second, if these properties are simply "picked out" by our concepts but are not shaped by them, one wonders what explains the coincidence between concepts and properties that is posited by PI. It is possible that we are simply lucky to have cognitive abilities so fortuitously formed, but it seems doubtful. What is more likely is that our concepts often, perhaps even usually, latch onto independently existing properties in the way suggested by PI, but not necessarily. If this is so, we should be on the lookout for signs that our concepts have not succeeded in picking out such properties.

In fact, there are signs that when it comes to qualia and appearance properties we are not latching onto independent, properly individuated properties. This becomes clear when we consider another reason to prefer the extensionalist modal of property individuation. We should expect properties with intensional individuation conditions to have trouble integrating into the causal and explanatory order of things. Assuming realism about the existence of properties, the intensionalist and the extensionalist should be committed to many of the same properties.³¹ After all, even the intensionalist must admit that in many cases our concepts are of properties that have completely objective individuation conditions—our concept of an electron, for example, or of mass. We should expect the intensionalist to have more properties in her ontology than the extensionalist, however, because in addition to those extensionally defined properties, there are some that can only be individuated partly in terms of the minds that know them. In the debate about conscious mentality, of course, we find that this is in fact the case: in addition to the physicalist's neural properties, the dualist has qualia. Thus the extensionalist's ontological commitments are a subset of the intensionalist's.³²

The result is that in contrast to the extensionalist, the intensionalist has what might be considered mixed domain of properties. The problem arises when it comes to the properties to which the intensionalist is uniquely committed—call these the intensional properties Assuming causation is an objective phenomenon—and it seems implausible that something's causing something else should simply be a matter of how we carve up the world—the complete causal story about the world should be able to be told in terms of properties that have objective individuation conditions.³³ Thus the extensionally individuated properties in the intensionalist's domain would seem to be

responsible for all of the causal action, leaving no causal role for the intensional properties.

Put another way, the intensionalist would seem destined to recognize properties that are not identical to the extensionalist's properties, and the latter are sufficient to explain the causal order of the world. If this is the case, then the intensionalist's properties are not causally necessary. Unless there is overdetermination in the case of intensional properties, which seems a slender reed for such an ontology, they are bound to be epiphenomenal. ³⁴ This is of course what tends to happen to qualia, but if my argument is right this has as much to do with the repercussions of the intensionalist method of property individuation as it does with any particular idiosyncrasies of conscious states.

If the foregoing is correct, we not only have a reason to prefer an extensionalist metaphysics to an intensional or mixed metaphysics, we also have our sign that in the case of appearance properties we are not latching onto extensionally individuated properties as the realist interpretation of PI would suggest. Recall that one could hold that PI was true not because properties are individuated by concepts, but because concepts pick out independently individuated properties. We can grant that such a coincidence between concepts and properties could obtain, but we should also alert to indications that it does not. We have found one sign: if properties lack extensional individuation conditions they will have difficulty integrating into the causal picture of the world. Since this is the case with the properties proffered by the dualist, we have reason to believe that at least in this case our concepts have not picked out independent properties but have led us to posit intensional properties with all of the problems that come with them.

This leaves us with the following situation. The presentation argument depends upon PI, which is itself committed to an intensional picture of the individuation of at least some properties. These properties are apt to be causally inert and, indeed, the properties found by the presentation argument are also apt to be epiphenomenal. This means that dualism gains plausibility because of a metaphysics that eventually drives it towards incoherence. On an extensionalist picture we can resist the presentation argument for dualism, adopting instead a coherent monistic ontology. This is the picture the subjective physicalist adopts, and it seems by far the preferable path.

The Conceivability Argument and the Metaphysics of Possible Worlds

In recent years, the presentation argument for property dualism has not been as popular as arguments from the conceivability of physical properties existing without mental properties.³⁵ According to these arguments, the conceivability of physical properties without mental properties indicates that it is possible that the two come apart. If it is possible that they come apart, they cannot be identical. David Chalmers has offered the most subtle defense of the conceivability argument in recent years, so I will focus on his presentation. Other versions of the argument can be handled in basically the same way.

In truth, the conceivability argument is a close sibling of the presentation argument, and it fails to undermine subjective physicalism for similar reasons. In particular, where the presentation argument presupposes a questionable metaphysics of properties, the conceivability argument presupposes a dubious metaphysics of possible worlds. The result, once again, is that at best it begs the question against subjective physicalism.

a. The Conceivability Argument

Chalmers summarizes the conceivability argument as follows:

According to this argument, it is conceivable that there be a system that is physically identical to a conscious being, but that lacks at least some of that being's conscious states. Such a system might be a zombie: a system that is physically identical to a conscious being but that lacks consciousness entirely. ...From the conceivability of zombies, proponents of the argument infer their metaphysical possibility. ...From here, it is inferred that consciousness must be nonphysical. If there is a metaphysically possible universe that is physically identical to ours but that lacks consciousness, then consciousness must be a further, nonphysical component of our universe. ³⁶

Subjective physicalism would seem to be particularly vulnerable to the conceivability argument. The subjective physicalist admits, after all, that there is an epistemic gap between physical descriptions and conscious states. Thus, it is conceivable that there be zombies with no "subjective properties," and this will be conceivable even given a completed objective science of the mind. Zombies therefore seem possible, which means—according to our own supervenience definition SVP—that physicalism is false.

The traditional response to conceivability arguments is to deny that conceivability entails possibility. This response is usually couched in terms of Kripkean *a posteriori* necessities: the fact that something is conceivable is ultimately an epistemic fact that does not inevitably reveal a metaphysical fact. We can conceive of the falsity of some necessities because they can only be discovered upon empirical investigation, but they are metaphysical necessities nonetheless. The conceivability argument has a response, however, as Chalmers—following in Kripke's own footsteps—has pointed out. When we are thinking of the falsity of an *a posteriori* necessity, we are envisioning a real possibility. Our mistake is to describe that possibility in such a way that it conflicts with the *a posteriori* necessity. This can be explained using what Chalmers calls "two-dimensional semantics."

According to two-dimensional semantics, the necessary *a posteriori* is best described as a phenomenon at the level of statements. "Water is H2O" is an example of the necessary *a*

posteriori, and this is because our language associates it with two different propositions or "intensions." The primary intension is the meaning of the statement gotten by considering a range of worlds as candidates for the actual one. It is that in virtue of which the statement picks out what it does in any world in which it is employed. The primary intension of "Water is H₂O", for example, might be captured by something like "the colorless, odorless stuff in lakes, rivers and oceans is H₂O." The secondary intension is the meaning of the statement taking the semantic facts of the actual world as fixed, and considering the statement counterfactually. In other words, it is what the Kripkean would consider the content of the statement, and the sense in which the statement is true in all worlds, with XYZ or otherwise. The primary intension is what determines whether a statement is a priori or a posteriori, and the secondary intension is what determines whether a statement is necessary or contingent. Statements that are necessary a posteriori, like "water is H2O," thus have a contingent primary intension but a necessary secondary intension.

Two-dimensionalism gains metaphysical bite when it is added that with the two intensions there are two senses of conceivability and possibility. To say that a statement is conceivable₁ is to say that one can conceive of a possible world where the primary intension is true, and to say that a statement is conceivable₂ is to say that one can conceive of a world where the secondary intension is true.³⁷ To say that a statement is possible₁ is to say that there is a possible world where the primary intension is true, and to say that it is possible₂ is to say that there is a possible world where the secondary intension is true. ("Possible" and "conceivable" are used in the *definientia* without subscripts because at the level of *propositions* conceivability and possibility are univocal. The subscripts just indicate which intensions are in the scope of the conceivability/possibility operators when they are applied to statements with two dimensions of meaning.) Given this, there no longer seems to be a problem with

inferring possibility from conceivability, as long as one only infers possibility₂ from conceivability₁ and possibility₁ from conceivability₁. (Though one can, of course, infer possibility₂ from conceivability₁ when the primary and secondary intensions are the same, as Chalmers claims is the case with thoughts about consciousness.) Concluding from the conceivability₁ of S that it is possible₂ is what the Kripkean thought experiments warn against: from the fact that we can conceive of its having turned out that watery stuff is not H2O we cannot conclude that there are worlds where water is not H2O. We can, however, conclude—by conceiving of the falsity of the primary intension—that there are worlds where watery stuff is not H2O.

If two-dimensionalism provides the real story underlying *a posteriori* identities—and it has a distinct air of plausibility when applied to the traditional Kripkean examples—then it should also apply in the case of mind-body identities. If pain is identical to a physical state f, then it is not possible₂ that there be a zombie world. But we are conceiving of something when we are conceiving of zombie worlds: we are conceiving₁ of a possible world—namely, a world where the primary intension of "pains are identical with physical state f" is false. To find this proposition we must locate the primary intension of "pain" and it seems that here the primary intension is something like "the unpleasant feeling that comes when I am wounded" and—assuming the *a posteriori* identity holds—the secondary intension is the basic physical description of "f." So, according to two-dimensionalism, we are licensed to infer that there is a world where "that unpleasant feeling" does not pick out anything despite the fact that there is brain state f. But this is a zombie world, since it involves the supposed physical part of pain without the feeling part. Thus, since two-dimensionalism vindicates the inference from conceivability₁ to possibility₁, the conceivability of zombie worlds shows their possibility, and by SVP physicalism is false.

b. The Conceivability Argument and Metaphysical Two-Dimensionalism

Two-dimensionalism is an impressive device because it seems prima facie intuitive while generating significant metaphysical results. One reason to be wary of two-dimensionalist arguments for dualism, however, is that what is intuitive about the device is distinct from what yields the metaphysical conclusions. For many of us, it seems obvious that there must be an element of thought content (if not of linguistic content) that captures the cognitive significance of our thoughts, and that this role cannot be satisfied by Millian or Russellian content. Thus, many of us are at least pre-theoretically committed to something like Chalmers' primary intensions, and Kripkean arguments usually persuade us that there are also secondary, Millian intensions. So far, however, this only commits us to a sort of minimal "cognitive" two-dimensionalism. This should not be confused with Chalmers' more "metaphysical" two-dimensionalism, however, and part of the persuasiveness of his arguments trades on our slipping from the former to the latter. Metaphysical two-dimensionalism requires that our primary intensions have their cognitive significance grounded in the existence of epistemically individuated possible worlds that answer to those intensions.³⁸ This is in contrast to what is perhaps a more natural view, that primary intensions are simply conceptual in nature, and they may or may not deliver possible worlds, which are individuated extensionally. It is the implicit ontology of metaphysical two-dimensionalism that provides the dualist results, and it is precisely that which we should question.

The intensionalist about possible worlds roughly maintains that there is a possible world corresponding to every set of circumstances that can be consistently thought. In particular an intensionalist about possible worlds will hold something like the following:

(PW) W is a possible world if it could be consistently conceived to be a complete state of affairs, and W is identical to possible world V only if it could be ascertained based on the complete conceptions of those worlds that V and W are identical.

In other words, every way one can conceive of a world as being corresponds to a possible world, and it is sufficient for difference between two possible worlds that they could not be identified based on *a priori* reflection.

An extensionalist, on the other hand, maintains that possibilities are what they are independently of what we can think about them, and *a priori* it is an open question whether our ability to conceive possibilities tracks the relevant modal facts. It's not necessarily the case that all worlds of which we can have a consistent conception are really possible, and not all possible worlds need be conceivable. Again, this is not to say that there are not inferential rules that connect conceivability to possibility, but such rules are merely defeasible epistemic licenses.

c. Against Metaphysical Two-Dimensionalism and World Intensionalism

For many, realizing that metaphysical two-dimensionalism is optional will be enough to lead them to reject the dualist argument. Dualism is, after all, an attractive haven only for those who are forced to occupy it. There are independent reasons to reject the world-intensionalism that is part of the dualist argument, however.

First, world intensionalism entails property intensionalism, so any skepticism about the latter should be cast upon the former.³⁹ This is especially obvious if one thinks of worlds as properties—"worlds" are really just ways the world could be.⁴⁰ In this case, the recognition of a world distinct from the class of extensionally defined worlds would

simply amount to the recognition of at least one intensionally defined property—the new world itself. One needn't view worlds in this way, however, for the entailment to hold. A contradiction can be derived by supposing PW is true and PI is false. For simplicity's sake, I'll consider the failure of only one direction of PI, namely that two properties are necessarily coextensive only if one can determine their coextensivity *a priori*. (The failure of the other direction of PI is also inconsistent with PW, but this is less crucial for my purposes.)

- 1. If PI is false, then there are some properties F and G that are necessarily coextensive, despite the fact that one cannot determine that coextensivity by reflecting on the concepts <F> and <G>. 41
- 2. For any two concepts <L> and <M>, if one cannot determine their necessary coextensivity by reflection, then one can conceive of a world w where there is an x such that Lx and ~Mx.
- 3. So if PI is false, one must be able to conceive of a world w where Fx and \sim Gx, despite the fact that F and G are necessarily coextensive.
- 4. So if PI is false and PW is true, it is possible that Fx and ~Gx, despite the fact that F and G are necessarily coextensive. This is a contradiction.

So if PW is true, PI has to be true as well.⁴²

Since world-intensionalism entails property intensionalism, the former inherits the profligacy, the coherence concerns, and the whiff of anti-realism that taints the latter. World intensionalism, as embodied by metaphysical two dimensionalism, also carries its own unique set of problems, however. In particular, it seems self-refuting.⁴³ Put roughly, the worry is that it seems conceivable that world intensionalism is false. According to two dimensionalism, if one can successfully conceive of the falsity of a statement, at the

very least the primary intension of that statement is possibly false. But since "World intensionalism is true" is presumably *a priori* and necessary (if true at all) its primary and secondary intensions coincide. So conceiving of its falsity does in fact deliver a world where world intensionalism is false. Since world intensionalism is necessarily true, if true, it's being possibly false means that it is actually false. If it is actually false then metaphysical two-dimensionalism is false, because without world-intensionalism there is no longer any guarantee that there is a world answering to our primary intensions. Thus, on the assumption of metaphysical two dimensionalism, metaphysical two-dimensionalism is false.⁴⁴

There is much more to be said about this *reductio* to two-dimensionalism than can be said here, but the general worry is clear enough. If possibility becomes tied too closely to conceivability, there is the risk that we can conceive of more things than can fit comfortably within a single logical space. In general, this is just a more pointed instance of the worry that individuating items epistemically makes it difficult to integrate those items into the objective order of things.

It might be objected that I have been unfair to the world-intensionalist. In particular, the more sophisticated defenders of the conceivability argument—including Chalmers himself—do not accept PW as stated. Instead, they accept a version of PW formulated in terms of *ideal* reasoners and conceivers, such as:

(PWI) W is a possible world if it could be consistently conceived to be a complete state of affairs by an ideal reasoner, and W is identical to possible world V only if an ideal reasoner could ascertain, based on the complete conceptions of those worlds, that V and W are identical.

PWI has the advantage that it does not tie possibility to the contingent abilities of

reasoners like ourselves, who presumably have limited conceptual repertoires and imperfect reasoning capacities. Nothing at all follows, for example, from the fact that someone thinking quickly can conceive of a possibility, since he might have overlooked some incoherence in his conception. Similarly, nothing follows from my ability to conceive something if I am an impaired reasoner or a poor conceiver. It is only ideal conception that counts.

This idealization move affects my previous arguments in two ways. First, it might seem to deflect the charge that the world-intensionalist is committed to a form of antirealism about possibilities. My charge might stick if possibilities were individuated in terms of the contingent abilities of fallible conceivers, but if possibilities are only tied to ideal conception it is substantially more plausible that such conceptions just pick out independently existing possibilities. The second effect of the idealization move is that my reduction might seem less plausible. Perhaps I can conceive that world-intensionalism is false—or that PWI is false—but that is no more telling my ability to conceive of the falsity of Fermat's theorem. I'm a poor mathematician and I haven't read Wiles' proof, so I can conceive of the theorem's being either true or false. Only conceivability by an ideal reasoner, someone much more like Andrew Wiles, would show anything about its possible (and in this case necessary) truth or falsity. The case seems analogous to my ability to conceive of the falsity of world-intensionalism. I might be a better philosopher than I am a mathematician, but an ideal reasoner I am not.

The idealization move is forceful, but I think it either winds up shifting the problem in a way that should significantly reduce the persuasiveness of the conceivability argument, or it does not go far enough to block my objections. The obvious question is how "ideal" conceivers and reasoners are being construed in PWI. A natural view is that

an ideal reasoner is one who doesn't make any mistakes. In this case, however, this seems tantamount to saying an ideal reasoner is one whose reasoning tracks possibilities. This won't do, of course, as it makes the conceivability argument blatantly circular: whether or not zombies are conceivable, for example, just amounts to whether or not they are conceivable by someone whose conceptions track possibilities. Since the conclusion is supposed to be that zombies are possible, the circularity of this reasoning is obvious.

There are other ways to analyze "ideal" in this context, and a survey of all the options deserves an investigation in its own right. It seems, however, that if the deflationary reading with its circularity is to be avoided, the notion of ideal reasoning needs to be thickened, perhaps by associating it with a type of reasoning that can be substantively defined, e.g. deduction. This can be done in one of two ways. It could be accomplished simply by reductively analyzing the notion of possibility in terms of that sort of reasoning. This route is independently implausible, but it risks either anti-realism about possibilities, or it begs the question by once again—though indirectly—identifying ideal reasoning as reasoning that gets possibilities right. The second way refuses to analyze possibility in terms of some system of reasoning, but only claims that the sort of reasoning in question is a perfect guide to possibility.

This last reading of PWI seems most palatable, but it still seems troubled. If possibilities are not reductively explained in terms of some manner of reasoning, we have to ask why that manner of reasoning is such a perfect guide to them. It's not clear what the explanation could be: infallibilism and realism seem unlikely bedfellows. This fact lends new blood to the *reductio*: assuming realism about possibilities, it seems conceivable that PWI is false, simply because if there is a domain of truths that is independent of a method of reasoning it is conceivable that the method of reasoning can occasionally miss

those truths. So what remains is a dilemma. Either two-dimensionalism is false by its own lights, or it is likely committed to a sort of anti-realism and threatens to introduce circularity into any argument that employs it.⁴⁵

It seems far preferable to say that conceivability is a good but defeasible guide to an extensionally individuated set of possibilities. When should we suspect that conception is defeated? Well, an obvious case is when it leads us into contradictions—as it might when applied to necessary truths such as PW or Fermat's Theorum. (In such cases, the conceivability of both the truth and the falsity of these propositions that are either necessarily true or false would generate a contradiction if the inference from conceivability to possibility were infallible.) A more subtle case of defeat is suggested by the discussion of the exclusion problem as it applies to property intensionalism. Suppose that a case of inferring possibility from conceivability leads one to posit a type of property that does not integrate well into the causal structure of the world. Suppose that one had to conclude that the "discovered" property lacked causal powers, such that it became mysterious how one knows about, refers to, or even thinks about such properties. These facts, I suggest, are signs that one's inference has been defeated.

Subjective Physicalism thus responds to the conceivability argument by embracing an extensionalist view of possible worlds. It is a corollary of this commitment that inferences from conceivability to possibility are defeasible. We should suspect that our conceptions have let us down when they commit us to properties that do not integrate into a coherent worldview and that seem fated to causal impotence. This, of course, is precisely what seems to be the case when it comes to property dualism. We should thus side with subjective physicalism, concluding that this is one instance when our trust in conceivability should bow to our commitment to a coherent worldview.

The Ontology of Subjective Physicalism

One can learn about everything in the world through objective methods, but there will nevertheless be something more to be learned by occupying some of the states studied by those objective methods. Those states are physical, but they cannot be fully grasped in the way we usually grasp physical states—they cannot be fully understood without undergoing them. This is not to say, however, that there are properties that are indescribable in the objective mode. Physics, as an objective approach to the world, might completely describe the world, in that it might be able to describe all of the properties of all of the things in the world. Some of those properties, however, are such that when they are possessed there is something it is like to possess them. In other words, while physics might be complete in describing all the properties in the world, there is a sense in which it does not describe all of those properties completely.

What, though, is this "something it is like?" Isn't it something that physics is leaving out? If Mary does come to know something she did not know before, mustn't that be a property that the world has that physical descriptions leave out? It seems we are back to the infelicitous choice presented by the dualist's dilemma. Either we must deny the reality of qualitative experience or we must let new properties, indescribable by physics, in the door.

There is a two-part response to this worry. One questions the demand for something ontological to explain the significance of Mary's knowledge, the other maintains that to the extent that something ontological is required, the ontology of subjective physicalism offers a unique way through what is left of the dualist's dilemma.

Underlying the dualist's dilemma is the sense that if Mary's epistemic gain does not reflect her discovering the ontological incompleteness of her previous view of the world it

path to the same old thing, then it seems her epistemic gain is no more significant than it would be if she learned the same old facts in Russian.⁴⁷ This argument assumes, however, that a subject's epistemic gain is only significant if it adds a new property or thing—considered extensionally—into her ken. This assumption is undermined—as is the Russian analogy—once it is realized that some ways of knowing things are clearly more valuable than others.

The nature and value of a subject's knowledge does not merely depend on what is known, but also upon the way that it is known. Knowing a physical state by experiencing it is a way of knowing that is of particular value to us. It is the way we first know about the physical states that underlie experiences, it is the way we most often identify those states, and it is in virtue of having that perspective on those states that we have the rich view of the world that we have. The subjective perspective is closely bound up with our conception of ourselves as agents and as thinking things. Although we might be able to conceive of the existence of zombies, we clearly cannot conceive of being zombies. There is a sense in which zombies have a perspective on themselves and on the world but we can hardly imagine that being our perspective. Lacking that perspective on an important class of things, therefore, would be a considerable deficit even if it is not a handicap that stems from an incomplete catalogue of the world's ontology. This alone explains, I think, why the perspective Mary lacks seems of particular importance. For all the greatness of the language of Pushkin, an inability to understand physical truths in Russian is simply not a comparable deficit.

One might agree that Mary's knowledge is more valuable than the knowledge of the same old things in Russian while still thinking that this does not really mark the importance of Mary's epistemic achievement. After exiting her room, Mary is in a position to rule out

ways the world might be that she was not able to rule out before.⁴⁹ This might suggest that she is actually coming to know about new properties.

This way of reviving the dualist's dilemma makes assumptions that subjective physicalism rejects. It illicitly presupposes an intensionalist view of properties, and if it assumes that the relevant "way the world might be" is a metaphysical possibility, it is assuming an intensionalist view of possible worlds as well. It is false to think that either Mary's "aha"-moment is purely psychological or she discovers a property that physics leaves out. In fact, Mary does not discover a property that physics leaves out, but she does learn about a property that allows her, given her cognitive economy, to rule out a set of "scenarios" that seemed to her consistent with what she had already learned. This has ontological implications only if one is committed to all such "scenarios" being possible worlds as opposed to merely conceptual possibilities, and the existence of these conceptual, epistemic possibilities only has ontological implications if we tie properties to concepts in an illicit way.⁵⁰

One way to describe what Mary learns, according to subjective physicalism, is to say that she comes to grasp an aspect of the property that she already knew about under its physical description. What, though, are these aspects? They are not themselves properties, but are instead part of the nature of properties that are not expressible by physical description.⁵¹ But how can they be part of the nature of properties without themselves being properties?

Consider the following analogy. According to the classical atomists, spatial atoms are the smallest parts of our world: they are the parts of our world that do not themselves have parts. One objection to this is that if these atoms are extended in space, then they have a front half and a back half. But if that is the case, then the halves seem to be parts that are

themselves smaller than the atoms, contrary to the atomistic hypothesis. The atomistic hypothesis thus seems contradictory. The solution for the atomists was to distinguish between an atom's actually being divisible versus its being conceptually divisible, and between its having real parts and its having conceptual parts. Real parts are parts that can be separated spatially from one another.⁵² Conceptual parts are parts that can be separated "in mind only": the mind can attend to them and distinguish between them, but in fact they constitute a basic physical unity. This does not mean that the spatial atoms don't really have front and back halves and that these halves are somehow in the mind. It just meant that these halves are not parts in a purely extensional, metaphysical sense.⁵³

Aspects are to properties as conceptual parts are to atoms. They are features of the properties that the mind can discern—and in the subjective case they are discernable only in virtue of possessing the relevant properties. They are not, however, separable from the whole of which they are a "part." In this case, separability obviously does not mean spatial separability. Instead, the subjective aspect cannot exist without the physical aspect and *vice versa*. This is one way in which aspects are distinct from properties. Another way they are different is that aspects are intensionally individuated while properties are not. As merely conceptual parts of the properties, they do not have individuation conditions that are purely metaphysical. This fits with the idea that "qualia" are somehow inextricably bound up with how they seem to a particular subject.

Subjective physicalism does not, therefore, recognize a distinct set of *properties* that correspond to qualia, unless those are taken to be the properties physics describes and Mary understands while still in her room. There are, to be sure, aspects of states and properties that are subjective, and these roughly correspond to qualia. It is in virtue of these aspects that Mary learns something when she leaves her room. But unlike qualia traditionally

conceived, the subjective aspects of certain physical properties do not enjoy independent metaphysical status. They cannot be a source of difference between objects, and they cannot become detached from the property of which they are aspects. They are not, therefore, prey to Churchland's *just more stuff* objection. To add these aspects to the list of physical properties would, in fact, be redundant—much as adding "the first half of atom A" and "the second half of atom A" to a list would be unnecessary, according to the atomist, if that list already included atom A.

Since aspects cannot be a source of difference between objects, and they are metaphysically dependent on the properties of which they are conceptual parts, they avoid certain problems that traditionally plague mental properties. In particular, aspects seem to sidestep the problems of mental causation. It seems that every event has a physical cause that is sufficient for it and thus that conscious mental properties make no causal difference. The intuition is that were there no such properties, the same effects would occur. This counterfactual gains its sense, however, from the separability of mental properties from physical properties. Since conscious aspects are not separable from physical properties, however, it is not the case that such aspects could be removed without causal difference: if those aspects were removed, so would the physical properties they are aspects of. If the conscious features of mentality were properties, this would not be the case. ⁵⁵

Subjective Physicalism in Contrast

As a penultimate note, it might be helpful to compare subjective physicalism to several nearby positions in the contemporary debate about conscious experience. In its broad brushstrokes it is similar to many of the positions on the table, but it is ultimately distinct in emphasis, motivation and in most cases it has fewer ontological commitments. In

general, what distinguishes subjective physicalism is precisely the emphasis upon the *subjective*; that is, upon the fact that for some states one must be in them to fully understand them. As such, subjective physicalism does not commit itself to any baroque apparatus or parade of metaphysical curlicues to explain conscious experience. Rather, it insists that we have all the makings of conscious experience given our physical constitution, but that this fact is not fully objectively explicable. Subjective experience is not something that can be grasped, much less explained, from outside the machine.

It might be thought that subjective physicalism bears some similarity to views that assimilate conscious knowledge to indexical knowledge.⁵⁶ Physics, or any other objective science, will notoriously lack descriptions in indexical terms, and this could be considered a sort of incompleteness.⁵⁷ What's more, it seems to be an incompleteness that derives from the objectivity of science: because it does not describe the world from a particular point of view, there will be nothing expressed in indexical terms. It appears theoretically satisfying, therefore, to say that the incompleteness of physics with respect to the subjective—and hence to conscious experience—is simply the same as the incompleteness with respect to indexical information. One can then give them both a fairly non-controversial semantic explanation that is completely consistent with physicalism.

Though subjective physicalism might seem to have a great deal in common with the indexical-knowledge views of consciousness, in the end the similarities are superficial. Indexical knowledge is no doubt closely linked with some sense of "subjectivity", but it isn't the sense adduced in the necessary condition for theory objectivity. A point of view or a location in the world is necessary for indexical utterances to be true, and the same can be said for beliefs with indexical components. It does not seem, however, that the existence of indexical knowledge requires that one occupy a particular state in order to fully understand

that state.⁵⁸ This is why we can suppose Mary has all the relevant indexical knowledge when she is in the room, yet still feel that she is missing something. She can think "I will see a rose at 3 pm today" and she can even point to a brain scan of someone seeing a rose and say "and when I see the rose I will have an experience like that." Nevertheless, she doesn't have the grasp of the state that she will later gain in virtue of occupying it. This and other criticisms of the indexical view have been well argued elsewhere, and they suggest a contrast with subjective physicalism.⁵⁹ Subjective physicalism requires, while indexicalism does not, that one must enter a determinate state of certain types in order to fully grasp states of that type.⁶⁰

Subjective physicalism might be more happily included among the various views Chalmers collects under the rubric "Type-B materialism" and that Frank Jackson calls "a posteriori physicalism." These views hold that while the facts about conscious experience are necessitated by the physical facts (i.e. SVP is true), they cannot be inferred a priori from those facts. Indeed, by this definition subjective physicalism is a form of Type-B materialism, but it is distinct and more satisfying than any of the views currently occupying that camp. Most such views employ phenomenal concepts to explain the particular sort of access we have to our conscious states. Accounts of phenomenal concepts themselves vary, but the basic idea is that such concepts employ physically explicable modes of presentation of conscious states that are not the same as the modes of presentation employed by concepts used in scientific categorization. It is furthermore at least implicitly presupposed that phenomenal concepts can be employed with respect to an experience only by the individual having that particular experience.

Subjective physicalism certainly allows that there could be phenomenal concepts that play an important role in introspection. Indeed, given that first-personal phenomenal

knowledge is distinct from its more objective counterpart, and given that this knowledge is portable—it can be retained beyond the occasion of the known experiences, and the concepts involved can be employed in a variety of thoughts—everyone should acknowledge that there are phenomenal concepts. Subjective physicalism does not, however, give the same explanatory role to phenomenal concepts as most Type-B materialist views. According to subjective physicalism it is simply a fact that there is something that it is like to instantiate certain physical states, and instantiating those states is sufficient for there being conscious experience—something that objective descriptions cannot fully capture. Since it is sufficient, phenomenal concepts are not necessary. The puzzling aspects of conscious experience are present before phenomenal concepts and discrete phenomenal beliefs enter the picture.⁶⁴ If this is wrong, and phenomenal concepts are necessary for conscious experience, it seems they are constitutive of that experience. In that case, they inherit the essential subjectivity of the experience and are not fully objectively explicable. So it seems that either phenomenal concepts are not explaining what needs explaining, or they are not themselves explicable.⁶⁵ Thus, although the subjective physicalist should acknowledge that phenomenal concepts exist, what is doing the work for the subjective physicalist is actually the more basic claim that there is something that it is like to instantiate certain physical states, and that it is only by instantiating those states that one can fully grasp them.

Ironically, in the end the subjective physicalist's closest friend is perhaps the property dualist. They both feel that there is something that it is like to instantiate certain states, that this can only be fully grasped by instantiating those states, and that objective sciences like physics leave this out. They differ, however, on whether physicalism can consistently acknowledge all three of these facts, and this difference has its source in the darker wilds of ontology. Subjective physicalism simply shows that the dualist's ontological commitments

upstream of his dualism are optional. To the extent that we feel that property dualism delivers an unlovely picture of our relationship to the world, threatening to make those most prized features of ourselves radically disconnected from the world at large, we should go back and choose metaphysical routes that allow us to embrace physicalism.

Conclusion

The mind-body problem is fueled by the sense that there is something special about conscious creatures despite the fact that they are constituted by the same stuff that constitutes the rest of the world. The former intuition seems undeniable, given that there is something that it is like to be a conscious creature. Denying the latter would seem to contravene the spirit of naturalism that seems increasingly irresistible as science progresses. We presumably evolved out of the same primordial swamps that gave rise to everything else, and nothing particularly unusual appears to have been added along the way. The problem is that we find it difficult to reconcile the continuity of constitution that holds between us and the rest of the world with a surprising discontinuity introduced by the presence of consciousness.

Subjective physicalism attempts to resolve this conflict of intuitions by insisting that we should not confuse our unique epistemic position with respect to some physical states with a metaphysical discontinuity. Metaphysically, we are of a piece with the rest of the world. We are, however, creatures that have states that can only be fully grasped by occupying those states. Many things in the world do not have such states. This is not to say that they are made of different types of stuff, however—that is a metaphysical claim. This is only to say that given the particular arrangement of physical stuff which constitutes them, there is no state that enables a particular sort of perspective on itself.

In this paper, I have aimed to defend *inclusive* subjective physicalism, according to which there are no properties left undescribed by physics even though physics does not provide the subjective grasp of some physical properties that is most important to us. In doing so, I have appealed to the existence of "aspects" of properties which are not themselves properties. No doubt there are those who will remain unconvinced that such aspects are not properties, despite the fact that they do not have the modal independence of properties and cannot be extensionally individuated. For such people *exclusive* subjective physicalism will appear a more attractive option—a complete physics would leave out "subjective properties," but such properties are physical nonetheless. Once the notion of property is expanded in this way, it must be admitted that there is little to distinguish between inclusive and exclusive subjective physicalism. At that point the principle complaint against inclusive physicalism would simply be that it embraces an intentionalist metaphysics of properties. In both positions, however the essence of subjective physicalism remains: everything is physical, and some physical properties enable an important subjective perspective on them.

Whichever version of subjective physicalism one accepts, it is clear that subjective physicalism does not put the mind body problem to rest, nor does it explain the subjective. It would be nice, for example, to have a full explanation as to why certain physical states have aspects that reveal themselves when those states are occupied. Subjective physicalism is not alone in its silence on this issue, however. Dualism itself cannot explain why only certain states are lawfully connected to the presence of qualia, and it doesn't seem that any view that offers a robust account of phenomenal knowledge can do much better.⁶⁷ No doubt the special sciences will eventually help us to some degree in this respect, but to some extent a complete explanation would require the elimination of any explanatory gap between

the subjective and the objective. Subjective physicalism denies that this is possible. We should not expect, because we cannot get, an objective explanation that captures all of the aspects of conscious experience. This is precisely what it means to say that experience is subjective. Thus there is still an explanatory gap, but it is perhaps less threatening since it simply falls out of the nature of objective theorizing. It might be no more mysterious than the fact that one cannot see dark matter or that one cannot analyze a poem using only mathematics.⁶⁸

References

- Alter, Torin. (1998). "A Limited Defense of The Knowledge Argument." *Philosophical Studies*, 90. pp. 35-56.
- Alter, T., and S. Walter. (2007) *Phenomenal Concepts and Phenomenal Knowledge*. Eds. Oxford: Oxford University Press.
- Block, Ned. (2007). "Max Black's Objection to Mind-Body Identity" in Alter and Walter (2007), pp. 249-306.
- Block, Ned, Owen Flanagan, and Guven Guzeldere, eds. (1997). *The Nature of Consciousness*.

 MIT Press: Cambridge.
- Campbell, Keith. (1990). Abstract Particulars. Oxford: Blackwell.
- Chalmers, David. (1996) The Conscious Mind. Oxford: Oxford University Press.
 - (1999) "Materialism and the Metaphysics of Modality." *Philosophy and Phenomenological* Research 59:473-93.
 - (2003). "Consciousness and its Place in Nature." In Stich and Warfield (2003).
 - (2004) "Imagination, Indexicality and Intensions." *Philosophy and Phenomenological* Research, Vol. XVIII, No.1.
 - (2006). "Phenomenal Concepts and the Explanatory Gap." Forthcoming in Alter and Walter (2006).
- Cohen, S. Marc, Patricia Curd and C.D.C. Reeve (2000) Readings in Ancient Greek Philosophy.

 Eds. "The Nichomachean Ethics" trans. Irwin, T. Indianapolis: Hackett.
- Crane, T. and D.H. Mellor (1990). "There is No Question of Physicalism." *Mind*, XC pp.185-206
- Crook, Seth and Carl Gillett. (2001). "Why Physics Alone Cannot Define the 'Physical". Canadian Journal of Philosophy. 31. pp.333-360.

- Dowell, Janice. (forthcoming). "The Physical: Empirical, not Metaphysical." *Philosophical Studies*.
- Ehring, Douglas. (1996). "Mental Causation, Determinables, and Property Instances." Nous. 30:4.
- Fantl, Jeremy and Robert Howell (2003). "Sensations, Swatches and Speckled Hens." *Pacific Philosophical Quarterly*. 84:4.
- Hawthorne, John. (2002) "Blocking Definitions of Materialism." *Philosophical Studies*. 110:2, pp. 103-113.
- Heil, John. (2003). From an Ontological Point of View. Oxford: Clarendon Press.
- Hill, C. S., and B. P. McLaughlin. (1999) 'There are Fewer Things in Reality Than Are Dreamt of in Chalmers's Philosophy', *Philosophy and Phenomenological Research* 59: 446-454
- Hill, C. S. (1997). 'Imaginability, Conceivability, Possibility and the Mind-Body Problem', Philosophical Studies 87: 61-85
- Jackson, Frank. (1982) "Epiphenomenal Qualia" Philosophical Quarterly XXXII, 127.(1986) "What Mary Didn't Know" The Journal of Philosophy LXXXIII, 5.(1998) From Metaphysics to Ethics. Oxford: Oxford University Press.
 - (2005). "The Case for A priori Physicalism." In Christian Nimtz and Ansgar Beckermann, eds., *Philosophy-Science -Scientific Philosophy, Main Lectures and Colloquia of GAP 5, Fifth International Congress of the Society for Analytical Philosophy*, 2003. Paderborn: Mentis, pp.251-265.
- Kim, Jaegwon. (1998). Mind in a Physical World. Cambridge: MIT Press.(2005). Physicalism, or Something Near Enough. Princeton: Princeton University Press.Levin, Janet. (2007). "What is a Phenomenal Concept?" in Alter and Walter (2007).

- Levine, Joseph. (2007). "Phenomenal Concepts and the Materialist Constraint." In Alter and Walter (2007).
- Lewis, David. (1999a). Papers in Metaphysics and Epistemology. Cambridge: Cambridge University Press.
 - (1999b) "New Work for a Theory of Universals" in Lewis (1999a) (1999c) "What Experience Teaches" in Lewis (1999a)
- Loar, Brian. (1997). "Phenomenal States" in Block, Flanagan and Guzeldere (1997)
- Mandik, Pete. (2001). "Mental Representation and the Subjectivity of Consciousness." *Philosophical Psychology*, Vol. 14, No.2. pp. 179-202.
- Martin, C.B. (1997). "On the Need for Properties" Synthesem 112: pp.193-231.
- McLaughlin, Brian. (2005). "A Priori versus A Posteriori Physicalism." In Christian Nimtz and Ansgar Beckermann, eds., *Philosophy-Science -Scientific Philosophy, Main Lectures and Colloquia of GAP 5, Fifth International Congress of the Society for Analytical Philosophy*, 2003. Paderborn: Mentis, pp.267-285.
- Melnyk, Andrew. (1997). "How to Keep the 'Physical' in Physicalism." Journal of Philosophy, Vol. 94, No. 12, pp.622-637.(2003). A Physicalist Manifesto. Cambridge: Cambridge University Press.
- Montero, Barbara. (1999). "The Body Problem." Nous. Vol. 33. No. 2., pp.183-200.
- Nagasawa, Yujin. (2002). "The Knowledge Argument Against Dualism." *Theoria*, 68. pp.205-233.
- Nagel, Thomas. (1979a) "What is it Like to be a Bat?" in *Mortal Questions*. Cambridge
 University Press: Cambridge.

 (1979b) "Subjective and Objective" in Mortal Questions. Cambridge University
 Press: Cambridge.

- (1986) The View from Nowhere. Oxford University Press: Oxford.
- Papineau, David. (2002). *Thinking About Consciousness*. Oxford: Oxford University Press. (2007). "Phenomenal and Perceptual Concepts." In Alter and Walter (2007).
- Papineau, David and Spurrett, D. (1999). "A Note on the Completeness of Physics."

 Analysis, 59.
- Perry, John. (2001). Knowledge, Possibility and Consciousness. Cambridge: MIT Press.
- Rodriguez-Pereyra, Gonzalo. (2002). Resemblance Nominalism. Oxford: Clarendon Press.
- Salmon, Nathan. (1986). Frege's Puzzle. Ridgeview Press.
 - (1989). "The Logic of What Might Have Been", The Philosophical Review, 98/1:3-34
- Schaffer, Jerome. (1961) "Could Mental States be Brain Processes?" *The Journal of Philosophy*, Vol.58, No.26, pp.813-822.
 - (1963) "Mental Events and the Brain" *The Journal of Philosophy*, Vol. 60, No.6, pp.160-166.
- Shoemaker, Sydney. (2003) "Causality and Properties." In *Identity, Cause and Mind*. Oxford: Oxford University Press. Pp.206-233.
- Siewert, Charles. (1998). The Significance of Consciousness. Princeton: Princeton University Press.
- Smart, J.J.C. (1959). "Sensations and Brain Processes." In David M. Rosenthal (1971)

 Materialism and the Mind Body Problem. Englewood Cliffs: Prentice Hall.
- Soames, Scott. (2005). Reference and Description. Princeton: Princeton University Press.
- Stalnaker, Robert. (1984) Inquiry. Cambridge: MIT Press.
 - (1999) Context and Content. Oxford: Oxford University Press.
 - (2003) Ways the World Might Be. New York: Oxford University Press.

Stitch, S. and Warfield, T. (2003). Blackwell Guide to the Philosophy of Mind. Eds. Oxford: Blackwell.

Sturgeon, Scott. (2000). Matters of Mind. London: Routledge.

White, Stephen. (forthcoming). "Why the Property Dualism Argument Won't Go Away."

(2007) "Property Dualism, Phenomenal Concepts and the Semantic Premise." In

Alter and Walter (2007) pp. 210-248.

Wilson, Jessica. (2005) "Supervenience Formulations of Physicalism" Nous. 39, pp. 426-459 (forthcoming). "On Characterizing the Physical" Philosophical Studies.

Yablo, S., (1999) 'Concepts and Consciousness', *Philosophy and Phenomenological Research* 59: 455-463

¹ I explore the possibilities of exclusive subjective physicalism in ***

² Such a commitment is the hallmark of the "Type-B physicalism" in Chalmers (2003). Authors such as Loar (1997), Hill (1997), Papineau (2002) and (2007), Sturgeon (2000), and many others all agree with this much. One might even include Lewis (1999c) though the cognitive achievement he has in mind is not the same as that suggested by these others or myself.

³ In fact I voice criticisms of most of those views in ***.

⁴ Nagel (1974), Jackson (1982) and (1986)

⁵ See ***

⁶ Alter (1998) also encourages this conclusion, as does Mandik (2001).

⁷ Nagasawa (2002) discusses Churchland's argument with a similar focus.

⁸ Churchland (1985) PP.24-5. While Jackson replies to this *tu quoque* argument in Jackson (1986), John Perry revives the criticism in Perry (2001). See section 7.5.

⁹ This is a slightly altered version of the principle I defend in ***. The alteration, which simply adds "or be able to enter," is meant to exclude missing-shade-of-blue-type interpolations from objective theorizing, as well as a class of other counterexamples presented and discussed by Torin Alter's "Phenomenal Knowledge without Experience" (forthcoming). Thanks to *** and *** for helpful debate about this principle. For a further defense, see ***. One objection is worth considering here, however. Objection: One must enter a state of performing reasoning, or of being awake in order to understand those states, even if they are understood using purest physics. Does that make any theory of them subjective? Answer: Being awake or being in a state of reason are not fully determinate states. One can be more or less awake, or can be reasoning more or less precisely, etc. The idea that one has to be able to enter into a particular fully determinate state of reasoning in order to understand that state doesn't seem plausible. I take the criteria for being a determinate/determinable property from Ehring (1996), who articulates and defends them in some detail. Thanks to an anonymous referee of this journal for raising this objection. ¹⁰ It is important to remember that subjectivity and objectivity are features of theories, points

of view or perspectives; they are not features of things or properties.

¹¹ See ***.

¹² This problem is well discussed in Montero (1999), Crane and Mellor (1990), and others. Analyses that might prematurely close the gap between the objective and the physical include Wilson (forthcoming) and Melnyk (2003).

¹³ I defend supervenience definitions against counterexamples such as necessitarian emergentist dualism, as presented in Wilson (2005) elsewhere. See ***.

¹⁴ This definition is inspired, in part, by one provided by Chalmers (1996) pp.39-40 and p.364.

- ¹⁸ It would be a mistake to think that because the negative definition defines the relevant notion of the physical in terms of the mental that the physical is not ontologically prior. It seems plausible, though, that in this debate the relevant notion of the physical is not *conceptually* prior to the notion of the mental, in that our grasp of the nature of the latter helps us refine our understanding of the former.
- ¹⁹ Two notes on this definition. First, I intend "something" in NIP in the broadest possible sense, ranging over objects, properties, events, etc. Second, something is fundamental in this sense iff it is a basic posit that is not reducible to another posit. Negative definitions such as this are used by Papineau and Spurrett (1999) and Papineu (2002).
- ²⁰ Subjective physicalism need not say that objective theories fail to provide a complete grasp of the world because they leave some properties undescribed. There is more on this—and the relationship between "aspects" and properties—in what is to come.
- ²¹ Smart (1959) p.63 The objection is apparently due to Max Black. See also Jerome Schaffer's (1961) and (1963), though in the end Schaffer seems more sympathetic to the type of view I propose than to simple property dualism.
- ²² Since the presentation problem does arise because of the existence of multiple epistemically individuated modes of presentation of one property, the role of mentality is not completely innocuous. It is not, of course, merely accidental that mental properties—such

¹⁵ Lewis (199b) is concerned with this problem, as are Jackson (1998) and Chalmers (1996).

¹⁶ As Jackson and Lewis do.

¹⁷ For more on this worry, see Hawthorne (2002).

should be suspicious when the model of reduction increases ontological commitments when mental properties are *not* the targets of reduction. There is a sign of a problem, in other words, before we ever attempt to reduce the appearance properties themselves. ²³ One of the only thorough critical discussions of the presentation problem I am aware of is Block (2007). Though the current paper was all but complete when I became aware of that piece, there are some similarities in our conclusions. Just how similar it is difficult to say, as our terminology and approach is different. Block does not focus on the individuation of properties, though it seems he is taking implicit stances on their individuation that are similar to those I support. Block thus does not criticize intentionalism as a metaphysics. Instead, he tries to dismantle arguments for some of the inferences intentionalism would sanction. ²⁴ Actually, this won't do, because an intensionalist might want to say that "equiangulartriangularity" and "equilateral-triangularity" are different concepts, yet they necessarily have the same extension. To get this result, one can distinguish between basic concepts and complex concepts, where the latter are concepts that are composed of other concepts and the former are not. Then one can say that concepts are distinct iff a) they are basic and a priori reflection cannot establish their necessary coextension, or b) they are complex and are composed of different basic concepts. Since equiangularity and equilaterality are not necessarily coextensive, and they are quite plausibly parts of "equiangular-triangularity" and "equilateral-triangularity," the latter can still be distinguished according to condition b.

as qualia—wind up being the ones most resistant to reduction. My point here is just that we

²⁵ Thanks to *** for helping me twist PI into shape.

²⁶ See Shoemaker (2003), for example.

The epistemic licenses would be defeasible and probably externalist principles, I suspect. It is my suspicion, actually, that the attraction to intensionalism is tied to an illicit attraction to a strong sort of internalist epistemology that would require, for example, knowledge of property identities to fall out of our concepts of those properties. I believe this epistemic internalism is implausible, but a defense of both that and my suspicion will have to wait.

Stephen White in his (forthcoming) presents an argument for intensionalism that appeals to its ability to get us out of Fregean puzzles. I both doubt its ability to get us out of those puzzles (see ***) and doubt that it is required to get us out of them (see Salmon(1996) for example). In addition, his argument to the contrary depends upon PW below, a premise the subjective physicalist will not let him have.

- ²⁹ Thanks to the referee for pushing this objection. Thanks also to *** who raised a similar issue.
- ³⁰ This is especially clear when it is argued that phenomenal properties ground the modes of presentation of informative identity statements about them. White (2007) is explicit about this, for example.
- ³¹ "Realism" here does not indicate a commitment to universals. It is to be contrasted with anti-realism, not nominalism.
- ³² An intensionalist might respond that it is inaccurate to view the extensionalist's domain as a subset of the intensionalist's. This might be so, but even so, an analogue version of this mental causation argument goes through—the general version stated later in the paragraph seems indifferent to the many ways one might count properties.
- ³³ It should be noted that I am here talking about the causal relevance of properties, not the appropriateness of making causal explanations employing reference to such properties.

Causal explanation might appropriately invoke properties that are not themselves causally relevant. This sort of thing is suggested by Yablo's proportionality constraint () and others. I am not inclined to think this mitigates the problem of causal relevance, but this debate deserves another locale.

³⁴ This is, of course, just a version of the argument from Kim, most easily presented in his (1998) and (2005). If what I have said is right, the problem of mental causation might be a species of a more general problem which doesn't stem from mentality so much as from the ways in which the relevant properties are individuated.

³⁵ Though Stephen White in his (forthcoming) does a good job of resurrecting the argument. He does so, though, in part by making it closer to the conceivability argument than it first appeared.

³⁶ Chalmers (2003) pp.5-6.

³⁷ See Chalmers (2002) for this way of putting things.

³⁸ Thus metaphysical two-dimensionalism is cognitive two dimensionalism plus a realist possible world semantics of thought. Both additions are optional.

³⁹ I actually think the entailment is mutual, but the other direction isn't particularly important for this dialectic.

⁴⁰ Stalnaker (2003). The entailment does not hold without supplementary assumptions even here—something like the assumption mentioned next is probably needed.

⁴¹ Note that as "F" and "G" are here names for properties, there is no assumption that F and G are distinct, so there is no implication here that denying PI forces one to accept the existence of necessarily coextensive properties. Thanks to ***.

 $^{^{\}rm 42}$ I think PI entails PW as well, though this is not required for my argument.

⁴³ I develop and defend this argument in *****.

⁴⁴ See...

⁴⁵ Supposing that the two-dimensionalist doesn't think possibility can be analyzed in terms of conceivability in some form, it seems as though he is committed to the coincidence between conception and possibility being a sort of "strong-necessity" of the sort two-dimensionalism abhors. This would significantly weaken Chalmers' objection to "type-B" materialisms to which Subjective Physicalism bears some resemblance.

⁴⁶ See Chalmers' own explanation of "The Paradox of Phenomenal Judgment" in Chalmers (1996).

⁴⁷ Lewis (1999c)

⁴⁸ See Siewert (1998), and ***.

⁴⁹ See Lewis (1999c), though he does not draw the ontological conclusions.

⁵⁰ Chalmers () talks in terms of such scenarios. These epistemic scenarios are in many ways the possible worlds analogues of the "aspects" of properties I introduce shortly.

[&]quot;Aspect" talk is not unheard of in this debate. See, for example, Block (2007) and White (2007). It is clear, however, that in the end they take aspects to basically be properties. Aspects of the sort I am talking about are not unheard of in the metaphysics of properties. John Heil and C.B. Martin seem to have something like this in mind when discussing the dispositional/categorical property distinction. I say this despite Heil's insistence that he is not talking about aspects, but I think this is because he presupposes that aspects of properties must be properties of properties. There is still a sense in which properties both have dispositional and categorical sides to them even if in some deeper sense they cannot be separated as "parts" of the property. See Martin (1997) and Heil (2003). Aspects of a

similar sort might also be necessary to make sense of the possibility of simple tropes that bear relations in a trope-theory such as that in Campbell (1990).

⁵⁴ If this is right, the aspect-theory part of subjective physicalism might be forced to deny that there are necessarily coextensive properties. This doesn't strike me as too much of a cost, and I would be willing to bite the bullet on this especially since such a view is independently supported by resemblance nominalism. See Rodriguez-Pereyra (2002) for this implication and for the rejection of some apparent counterexamples.

⁵² The modality here is metaphysical, not nomological.

⁵³ Cf. Aristotle when he asks of two parts of the soul whether they are "two only in account, and inseparable by nature, as the convex and the concave are on a surface?" Cohen, Curd and Reeve (2000) p.775 Other potential examples of aspects are equiangularity and equilaterality, colors and shapes, etc.

⁵⁵ For further discussion of this issue see my ***. See also Bennett ().

⁵⁶ See, for example, Perry (2001).

⁵⁷ See Perry (1979) and Lewis (1979).

 $^{^{58}}$ This point is made nicely by both Jackson (1986) and Mandik (2001).

⁵⁹ For criticisms, see Chalmers (2004), Block (forthcoming) and ***

⁶⁰ It might be the case that an explanation of the full cognitive significance of indexicals must make reference to states that are subjective in my sense—as I argue in ...--but if that is the case, the epistemic uniqueness of conscious states explains the uniqueness of indexical knowledge, not the other way around. This is, I think, the intuitive result.

⁶¹ See Chalmers (1999) for example, and Jackson (2005).

⁶⁴ This suggests a view to which I am independently attracted, namely that there is a type of privileged access we have to our conscious states that is not fully reflected in our beliefs about those states. This is a candidate for a sort of non-conceptual knowledge. I defend this view in ***. One needn't accept this, however. The point here is not necessarily that there can be non-conceptual consciousness, it is just that the burden of explaining consciousness cannot be borne solely by phenomenal concepts. Thanks to *** and *** for pressing me on this point.

⁶² Loar (1997) is often seen as the originator of this strategy, but others, including Papineau (2002) follow him.

⁶³ It is not always obvious why this is the case, however. If one were able, through sophisticated neural wiring, to apply one's phenomenal concept to someone else's experience, would one come to know what it's like to have that experience without actually having it?

⁶⁵ On this count, subjective physicalism should agree with the dilemma that Chalmers (2006) poses for the phenomenal concept strategy.

⁶⁶ I defend a version of exclusive subjective physicalism in ...

⁶⁷ Phenomenal Concepts accounts would be superior in this respect, were they not subject to the criticisms I level against them in the previous section.

⁶⁸ Acknowledgments...