

J.L. DOWELL

## FORMULATING THE THESIS OF PHYSICALISM: AN INTRODUCTION<sup>1</sup>

Perhaps more controversial than whether physicalism is true is what exactly would have to be true for physicalism to be true. Everyone agrees that, intuitively at least, physicalism is the thesis that there is nothing over and above the physical. The disagreements arise in how to get beyond this intuitive formulation. Until about ten years ago, participants in this debate were concerned primarily with answering two questions. First, what is it for a property, kind, relation, or individual to be a physical one? Answers to this first question offer accounts of what it is to be physical in a primary sense. Second, what is it for a property, kind, relation, or individual to be nothing over and above the physical ones? Answers to this second question offer rules for extending the class of physical properties, etc. beyond the basic, primary ones. I'll call these properties 'physical in a secondary sense'.<sup>2</sup>

Within the last ten years, many philosophers of mind have also begun to think about a third issue concerning what would follow from physicalism's truth and so what else would have to be true for physicalism to be true. Very roughly, this debate is over whether physicalism's truth requires that there be *a priori* entailments from the physical truths<sup>3</sup> to all of the truths. Each of the papers contained in this volume addresses one of these three issues. In this introduction, I propose to summarize the positions defended in each of these papers as well as to situate those positions in the literature as a whole.

### 1. CHARACTERIZING 'THE PHYSICAL'

The first three papers in this volume defend different answers to the question concerning what it takes for something to be

physical in the primary sense. In each case, the aim is to develop an account that adequately addresses skeptical concerns about whether a notion of the physical can be developed that is not circular, possesses ‘sufficiently determinate’ content, and allows for a genuine question as to physicalism’s truth. A further common aim is to maintain adequate overlap with both our pretheoretical concept of the physical and with the notion as it has figured in the literature on the mind–body problem.

Behind many of these concerns with the notion of the physical is what has come to be known as Hempel’s dilemma. The dilemma arises for accounts that aim to flesh out that notion by appeal to physics. There are two ways an account might appeal to physics. On the first strategy, the properties, kinds, relations, and individuals identified as physical in the primary sense are those posited by our current physics, while the second identifies the primary, physical properties by appeal to the posits of our ideal physics. The dilemma is that each of these strategies results in a puzzle. Induction over the history of physics tells us that future physics will posit properties, relations, or kinds that current physics does not posit. Further, it tells us that this future physics will replace current physics by being more warranted by evidence gathered in accordance with the very standards we treat as warranting current physics, that is, broadly speaking, by its providing a better explanation of the observations it is the business of physics to explain. So where current and future physics differ in their posits, we will have better grounds to believe in the existence of what future physics posits. In short, this induction tells us that we already have grounds for thinking that current physics is false, at least by being incomplete.

Consider now how appeal to current physics would fill out ‘the physical’ part of our intuitive formulation of physicalism as ‘nothing over and above the physical’. Filling out, the resulting thesis would be that there is nothing over and above the posits of current physics. But our simple induction tells us that we already have good grounds for thinking that physicalism, on this formulation, is false. Given that the

posits of physics are relatively basic, on this formulation, when future physics adds to the stock of current physics' posits, it is positing the existence of something that is not physical. Since our simple induction tells us both that future physics will add to the posits of current physics and that it will be more warranted than current physics by a common standard, we already have good grounds for thinking that there is some such posit. And that there is makes physicalism on this formulation false.

What is the lesson of this horn of Hempel's dilemma? The lesson is not, of course, that physicalism is false, but that a formulation of physicalism in terms of current physics is inadequate. Isolating where the simple induction shows that formulation to be inadequate teaches us that there is a constraint any formulation must meet in order to be adequate. What is inadequate about current physics formulations is that if we accept them, we seem to know something that we do not in fact know, namely, that physicalism is false. Even if physicalism on its current physics formulation is false, there still seems to be a question as to whether there is anything over and above the physical. We might capture this thought in what I call 'the Genuine Question Constraint'.<sup>4</sup> Meeting that constraint requires that a formulation of physicalism allow for a genuine question as to its truth.

On the second strategy, the physical is characterized in terms of the posits of future, possibly ideal, physics. On most readings of Hempel's dilemma, the difficulty with this strategy is that either it threatens to deprive physicalism of determinate content or the content it delivers renders physicalism trivially true.<sup>5</sup> The first of these difficulties is generated from the incontrovertible observation that we don't know what, exactly, future physics will posit. Given this, the argument goes, any formulation of physicalism that results from an appeal to future physics is no more than hand-waving; the thesis is no more determinate than "there is nothing over and above *that*". The difficulty, then, is that if the content of physicalism is so vague and indeterminate, its truth-value

becomes impossible to evaluate. The problem is not just that we can't now evaluate physicalism's truth. The problem is that we can't now say anything about what would and wouldn't be compatible with its truth. Given that physicalism is an *a posteriori* thesis, it should have a content such that it is clear which of the ways the actual world might turn out to be are ways that would make physicalism false at the actual world. Meeting this objection involves satisfying what I call the 'the Content Constraint', the requirement that a formulation possess enough content so as to allow for the identification of what would be its turning out that physicalism, on that formulation, is false.<sup>6</sup>

The second difficulty many see arising on this horn of Hempel's dilemma is that future or ideal physics characterizations of the physical threaten to make physicalism trivially true. To generate this concern, one needs to assume that future physics will be a theory of everything or, at least, a theory of every contingent or *a posteriori* truth. The claim is that whatever is found to be not reducible to the posits of physics will ultimately be incorporated into physics. But if this is so, there will be no posit of future physics that could falsify physicalism formulated as a thesis about the contingent or *a posteriori* truths about our world. So physicalism is rendered trivially true.<sup>7</sup>

It is a bit hard to see why one should think that future physics will be all-encompassing in the way needed to generate this objection. Further, I argue that if we think about what we now know about what kinds of properties couldn't be posits of a genuinely physical theory, it is not so hard to give physicalism defined in terms of ideal physics enough content so as to be able to identify ways the world could turn out such that physicalism would be false.<sup>8</sup> Partly for these reasons, Jessica Wilson offers a third account of the central difficulty faced by ideal physics accounts.<sup>9</sup> According to Wilson, the central difficulty such accounts face is that nothing rules out in principle that among the posits of ideal physics will be properties that are "intuitively physically unacceptable...most problematically...entities that are fundamentally mental."<sup>10</sup> According to

Wilson, there is no way for fundamental, mental properties to be compatible with physicalism's truth. On these grounds, she argues that a characterization of the physical in terms of the posits of ideal physics must add a 'no fundamental mentality' (NFM) constraint. The account she defends is

*The physics-based NFM account:* An entity existing at a world *w* is physical if and only if

- (i) it is treated, approximately accurately, by current or future (in the limit of inquiry, ideal) versions of fundamental physics at *w* and
- (ii) it is not fundamentally mental (that is, does not individually either possess or bestow mentality).<sup>11</sup>

In contrast, in my paper "The Physical: Empirical, not Metaphysical" I defend an unadorned ideal physics-based account. What separates the two accounts is the acceptance or rejection of the NFM constraint. So deciding between them turns on whether there are good grounds for its imposition. Wilson's grounds for its imposition are (1) there is a widely accepted intuition that any fundamental, mental property would falsify physicalism, (2) its imposition is required for physicalism to serve as a recognizable ancestor to materialism, (3) its imposition is required in order to so much as state the traditional mind-body problem, and (4) its imposition is required to preserve the contrasts between physicalism, on the one hand, and emergentism, epiphenomenalism, and substance dualism, on the other.<sup>12</sup>

I accept #1 as datum, but offer a deflationary explanation of the intuition. Further, I argue that imposing the NFM is not necessary for an account to possess all of the features Wilson rightly identifies as desiderata in #2–#4.<sup>13</sup> The crux of the issue, it seems to me, is whether or not one holds that the incompatibility between fundamental mentality and physicality is *a priori* or *a posteriori*. In my reserved (and non-idiosyncratic)<sup>14</sup> sense,

a claim *c* is held *a priori*...just in case there is no way the world could turn out to be such that *c* is false. To be held *a posteriori* requires holding that there is *some* way, however extremely unlikely, the world could turn out to be such that *c* is false.

Whether or not one holds that the incompatibility between fundamental mentality and physicality is *a priori* or *a posteriori* in these senses turns on whether or not one recognizes a way the world might turn out to be, however remote, such that they are compatible. To impose the NFM constraint is in effect to refuse to recognize anything as a way the world could turn out to be such that something (e.g. a property) is relatively fundamental, physical, and mental. So one who accepts the NFM constraint holds the incompatibility between fundamental physicality and mentality as an *a priori* matter, in the present sense. Given this, defense of that constraint requires some reason to think that the incompatibility is *a priori*.<sup>15</sup>

On the other side, the plausibility of a view that treats the incompatibility as *a posteriori* requires identifying a way the world could turn out to be that is plausibly described as its turning out that the two are compatible, given what is central to our concept of the physical for the purposes of formulating physicalism. On my unadorned, ideal physics-based account, physical properties, relations, and kinds are those posited by the complete and ideal scientific theory of the world's relatively fundamental elements. This simple formulation is supplemented by accounts of 'complete', 'ideal', and 'scientific theory'. The result is a strict requirement on what is to count as a basic physical property, relation, or kind.

To count as basic and physical, a property must be well integrated into the most complete and unified explanation possible for the relatively basic occupants of space–time. To be so integrated, its behavior must be highly regular. So it is not enough...that a property's instantiations are merely compatible with events explained by the ideal and complete physical theory. It must be well integrated into its overall pattern of explanation.<sup>16</sup>

In holding the incompatibility of the basic, physical and the mental *a posteriori*, I hold that it would be extremely surprising, though not epistemically impossible, for there to be mental properties that possessed the features required to be basic, physical ones in this sense. Since it is close to incontrovertible that a central part of our notion of the physical as

it figures in physicalism is tied to what physics tells us there is, this scenario is very plausibly described as one in which an extremely remote epistemic possibility is shown to be actual. This allows that the unadorned, ideal physics-based view I defend to treat the incompatibility as *a posteriori* while retaining its own plausibility.

Sara Worley takes a different tack in her paper “Physicalism and the Via Negativa”. Her central aim is to undermine Barbara Montero’s claim that the notion of the physical cannot be given sufficient content to figure in an adequate formulation of physicalism.<sup>17</sup> Montero’s strategy is to give up the project of defining ‘the physical’ as it figures in physicalism and replace the question as to physicalism’s truth with the question as to whether there are fundamental, mental properties. Worley aims to defend the claim that we do have a concept of the physical with determinate content, but to argue that that content is partly constituted by “a contrast with the mental”.<sup>18</sup>

She appeals to empirical psychology to show that

the distinction between the mental and physical is hard-wired...more precisely...we have, as part of our basic biological endowment, a propensity to develop two different systems or modes of explanation of behavior: one for agents and one for (mere) physical objects.

The literature in developmental psychology suggests that at an early age, children develop a proto-theory of the physical according to which what’s physical is “cohesive, solid,...move(s) continuously (i.e. ...traces a continuous path through space rather than ‘jumping’ from one spot to another), and...move(s) only via contact”. Objects children interpret as possessing mentality are regarded, in contrast, as “self-propelled”.

Worley recognizes that the first three of these elements in a child’s proto-theory of the physical are not useful for the purposes of giving an account of the notion of the physical as it figures in physicalism. The reason is that there are entities, kinds, properties, and forces that we recognize as physical on the grounds that our best physical theory tells us they are, yet

fail to have one or more of these three features. Worley concludes that it is the final, fourth feature, self-propulsion, that distinguishes the mental from the physical. Mental properties allow for the self-propulsion of what possesses them, while physical properties require external forces for their causal powers to be activated.

One challenge for Worley's account is to explain why this aspect of our proto-theory is central for the purposes of identifying our concept of the physical as it figures in the philosophical formulation of physicalism. A perhaps deeper problem, however, is that the account appears *ad hoc*. If being a part of our proto-theory of the physical is what makes something properly a part of our concept as it figures in a formulation of physicalism, then there is no good grounds for excluding the first three features of our proto-theory, except, of course, that including them fails to capture our widely-held intuitions about the extension of our concept of the physical. But the aim in giving an account of our concept as it figures in a formulation of physicalism is to say what something needs to be like in order to count as physical, not merely to get the concept's extension right.

What good grounds are there for not including solidity, cohesiveness, and continuity of movement as relevant parts of our concept of the physical despite their being part of our proto-theory? The answer, of course, is failure to fit with our best physical theories. But if an incapacity for self-propulsion is a physical-making feature on the grounds that it fails to conflict with those theories, then what really justifies Worley's account of our notion of the physical is its failure to contradict physics, not that it is part of the proto-theory nature seems to endow us with. If the adequacy of an account of the aspects our concept relevant for formulating physicalism stands or falls based on its fit with physics, then it is our sense of what physics will tell us there is that underlies our concept of the physical as it figures in physicalism, not our proto-concept, for all its use in explaining the behavior of children.



## 2. CHARACTERIZING 'NOTHING OVER AND ABOVE'

Once the physical in its primary sense has been characterized, physicalists need a physicalist-friendly extension rule, that is, a rule that says how properties that are not primarily physical must be related to the primary ones compatible with being physical, albeit not primarily so. Philosophers have considered a variety of extension rules in this sense, e.g. token-event identity,<sup>19</sup> type-type identity,<sup>20</sup> supervenience,<sup>21</sup> and physical realization.<sup>22</sup> In the first of two papers addressing this topic here, Andrew Melnyk defends a formulation of realization physicalism against rival realization-based accounts.

Realization-based accounts are attractive because they are able to accommodate what might at first blush seem to pose a difficulty for physicalists, namely, the multiple-realizability of some higher-order states and properties, mental ones, in particular. A higher-order property is multiply realizable if and only if it is possible for two entities to differ in their lower-order properties while sharing a single higher-order one. The trick for the physicalist is to identify the relation between the higher-order property and lower-order ones in such cases such that the higher-order property is nothing over and above the lower-order ones.

The inspiration for realization-based accounts is the relationship between a computer's hardware and its programs. It is clear that two computers can differ in their hardware, while sharing a capacity to run the same programs. Moreover, in such cases, the program does not seem to be something extra, over and above each computer's hardware. The reason for this is that programs are functions. Two objects may be physically different and yet able to perform the same function. When they do, the function performed does not seem to be something over and above the physical properties and relations that allow make performance possible.

Realization-based formulations of physicalism in Melnyk's sense are characterized by their fleshing out of the following schema

(RBFP) Everything that exists is either *identical with* the physical (in a certain narrow sense of ‘physical’) or *realized by* the physical (in the same narrow sense of ‘physical’).<sup>23</sup>

Realization-based formulations diverge in their accounts of the realization relation. Part of what makes Melnyk’s account so plausible is that it seems to explain what’s behind our intuition that a computer’s programs are nothing over and above the properties of its hardware. It does so by building into its constraints on token realization a condition that requires that the realized property is a functional one. On that account,

(RP-R) Token *x* *realizes* token *y* (or: token *y* is *realized by* token *x*) iff  
 (i) *y* is a token of some *functional* type *F* (i.e. some type whose tokening just is the tokening of some or other type that meets a certain condition, *C*);  
 (ii) *x* is a token of some type that in fact meets *C*; and  
 (iii) the token of *F* whose existence is necessitated (in the strongest sense) by the holding of clause (ii) is numerically identical with *y*.<sup>24</sup>

To get the realization-based formulation of physicalism Melnyk defends, we first define physical realization using the above account of realization to get

(RP-PR) A token *y* of a functional type, *F*, is physically realized iff  
 (i) *y* is realized, in the sense of RP-R, by a token of some physical type, *T*; and  
 (ii) *T* meets the special associated condition for *F* solely as a logical consequence of (a) the distribution in the world of physical tokens and (b) the holding of the physical laws.

The RBFP that results is

(RP) Every causal or contingent token of any type – whether an object, property-instance, or event – is either (1) a token of a physical type or (2) a physically realized token of a functional type.

Melnik’s aim is to consider how (RP), a view he defends at length elsewhere,<sup>25</sup> fares against three rival formulations that result from filling in the (RBFP) schema with three different accounts of realization found in the literature on the topic. The first formulation Melnyk considers is developed from the Ernest LePore and Barry Loewer account of realization,<sup>26</sup> while the second and third formulations are elaborations of

two accounts of realization found in the work of Sydney Shoemaker.

Melnyk's strategy is first to identify a set of conditions any formulation of physicalism must meet in order to count as genuinely physicalist. The first such condition is a constitution condition; in order to be genuinely physicalist, a formulation must identify a sense in which all phenomena are constituted by the primarily physical phenomena. The second is a truth-making condition. This condition requires that a formulation specify how it is that all truths are made true by what makes the primary, physical truths true. Finally, a formulation must meet a third, necessitation condition. In order to be genuinely physicalist, a formulation must have as a consequence that for any object one takes, its physical properties, relations, and the laws its properties figure in together "necessitate in the strongest sense" all of its properties. Melnyk then argues that (RP) more plausibly satisfies these conditions than the rival realization-based formulations he considers.

The first part of Terence Horgan's paper "Materialism: Matters of Definition, Defense, and Deconstruction" also addresses the issue of how best to understand the relationship between the primarily physical properties and all of the properties needed for physicalism's truth. Here, Horgan defends a supervenience-based formulation which may be thought of as a modification of Frank Jackson's supervenience-based formulation. According to Jackson, the following supervenience thesis both entails and is entailed by physicalism:

(MPD) Any world which is a minimal physical duplicate of our world is a duplicate *simpliciter* of our world <sup>27</sup>

Following Witmer<sup>28</sup>, I'll call this thesis the 'minimal physical duplicate thesis'.

A *minimal physical duplicate* of our world is a world that is identical to our own in its instantiation and distribution of physical properties, kinds, and relations and contains nothing else. MPD says that all such minimal duplicate worlds are exact duplicates of the actual world. It is easy to see that if

physicalism is true, then so is MPD. To see this, suppose that MPD is false. Then minimal physical duplicate worlds lack some qualitative feature *F* of the actual world. But since minimal physical duplicate worlds are exactly like the actual world in all of their physical respects, *F* must be non-physical. But then the actual world has a non-physical feature, namely, *F*. So if MPD is false, then so is physicalism. Hence, if physical is true, then so is MPD.

But the reverse entailment fails if there is no incoherence in the possibility of metaphysically necessary connections between distinct existences one of which is physical and the other, non-physical. To see this, suppose that physicalism is false. If physicalism is false, then the actual world contains a non-physical feature, *N*. If *N* is necessarily connected to some physical feature or features, *P*, then any world that duplicates all of the actual world's physical features will duplicate *P* and hence, also *N*. So if there is no incoherence in the possibility of metaphysically necessary connections between *N* and *P*, then MPD may be true while physicalism is false.

Several philosophers, including Horgan here, have argued for just such a failure of incoherence.<sup>29</sup> If they are right, then (MPD) is not equivalent to physicalism. Horgan's formulation here may perhaps be best thought of as a revision on Jackson's basic proposal so as to accommodate just this possible coherence. It is:

- (M3) (1) The actual world is a minimal physical duplicate of itself,  
 (2) for any physically possible world *w*<sub>1</sub> and *w*<sub>2</sub>, if (i) *w*<sub>1</sub>\* is a minimal physical duplicate of *w*<sub>1</sub>, (ii) *w*<sub>2</sub>\* is a minimal physical duplicate of *w*<sub>2</sub>, (iii) *r* is a spatiotemporal region of *w*<sub>1</sub>\*, (iv) *s* is a spatiotemporal region of *w*<sub>2</sub>\*, and (v) *r* and *s* are intrinsically just alike in all physical respects, then *r* and *s* are just alike in all intrinsic respects,  
 and  
 (3) there are no brute inter-level relations of metaphysical necessitation linking physical particulars or properties to non-physical particulars or properties.

The remainder of Horgan's paper is devoted to the issue of what is required to establish physicalism's truth, an issue I discuss in the next section.

## 3. A PRIORI VERSUS A POSTERIORI PHYSICALISM

During the last ten years or so, a debate has emerged about whether or not physicalists are committed to the *a priori* truth of certain conditional statements.

It is easy to show that if MPD is true, then so is a thesis that Gene Witmer calls “the physicalist entailment thesis” or:

(PE) Necessarily, if PT, then A.<sup>30</sup>

In (PE), ‘P’ is the lengthy conjunction of all of the physical truths at the actual world and ‘T’ a ‘that’s all’ clause, a clause that says of what makes ‘P’ true that there is it and nothing else. (Thus, ‘PT’ will be true at all and only the minimal physical duplicates of the actual world.) ‘A’ is a sentence that gives the true, complete description of the actual world (or at least of all contingent truths). Finally, ‘necessarily’ here is used in the medium-strength sense of ‘necessarily truth-preserving’, that is, there are no worlds at which ‘PT’ is true that are not also worlds at which ‘A’ is true.<sup>31</sup> So interpreted, (PE) is a neutral way of capturing the physicalist claim that all of the truths are necessarily determined by the physical ones, accepted by both *a priori* and *a posteriori* physicalists.<sup>32</sup>

The dispute between *a priori* and *a posteriori* physicalists most discussed in the literature is over whether or not physicalists are committed to holding that (PE), or perhaps a non-modal version of (PE), is *a priori*.<sup>33</sup> *A priori* physicalists are so-named for holding that they are, *a posteriori* physicalists for holding that they are not. On Gene Witmer’s reconstruction of the issue, strict *a priori* physicalists are characterized by their acceptance of two additional assumptions needed to take one from (PE) to that (PE) (or a non-modal counterpart of (PE)) must be knowable *a priori*. The first is a thesis about the knowability of necessities that he calls ‘the *a priori* accessibility of necessity thesis’ or APAN:

(APAN) If it is necessary that P and it is knowable that it is necessary that P, then either (i) it is *a priori* knowable that it is necessary that P; or (ii) there is some true non-modal sentence ‘E’ such that it is empirically

knowable that E and it is *a priori* knowable that if E, then it is necessary that P.<sup>34</sup>

So, given APAN, the knowable necessity of PE means that either PE is itself *a priori* or a conditional that has some E as its antecedent and PE as its consequent is *a priori*. If we in addition assume with the strict *a priori* physicalists that whatever information 'E' contains is knowably contained in PT (as seems plausible if MPD is true), then

(L) If E, then necessarily, if PT, then A

can be simplified to

(S) If PT, then A<sup>35</sup>

the non-modal version of PE. So the non-modal version of PE is, as the strict *a priori* physicalist claims, itself *a priori*.

According to Witmer, the rock on which this argument founders is the mistaken second assumption that the information contained in 'E' is *a priori* recoverable from PT, an assumption he calls 'the physical information assumption'. Witmer's own view is motivated centrally by the rejection of this assumption. His key idea is that the rejection of the physical information assumption does not mean the rejection of *a priori* physicalism, only its modification. The modified view he calls 'liberal *a priori* physicalism'. The contrast between the strict *a priori* view, defended by Frank Jackson, and the liberal view is this:

The strict *a priori* physicalist accepts the physical information assumption and...the argument for thinking that it is *a priori* knowable that if PT, then A. The liberal *a priori* physicalist is not committed to the *a priori* knowability of that truth; he is only committed to a weaker thesis, namely, that there is some true statement 'E' about the actual world such that it is *a priori* knowable that if E, then necessarily, if PT, then A.

The liberal *a priori* physicalist is, then, quite liberal in one very important respect: he allows that the formulation of what is *a priori* may require the use of terms from outside of physics in the initial antecedent.<sup>36</sup>

In the remainder of his paper, Witmer defends a revision of APAN so as to exclude a reading that would render it trivi-

ally true and so not a source of contention with *a posteriori* physicalists. He then outlines a strategy for the liberal *a priori* physicalist to show how zombies are not metaphysically possible and so to block one threat to physicalism's truth. Zombies would be creatures physically identical to ourselves, but lacking consciousness. If zombies are metaphysically possible, then MPD is false. But if MPD is false, then, as we've seen, so is physicalism. So, in order to rebut the dualist's challenge, physicalists must provide some grounds for thinking that zombies are not possible.

There is much disagreement (some of it discussed below) about what is required to provide such grounds. As we've seen, both *a priori* and *a posteriori* physicalists accept PE. If the truths about consciousness are entailed (in the necessarily truth-preserving sense) by PT, then zombies are not metaphysically possible. One disagreement between *a priori* and *a posteriori* physicalists is over what is needed to show that there is such an entailment. Here Witmer argues that it would be enough to show that there is an *a priori* conditional that has as its antecedent some true sentence that says of some physical state P that it plays the role of some phenomenal state S, where the relevant phenomenal role is itself specified using psychological vocabulary, as is permitted on the liberal *a priori* view. The consequent would then say 'S=P', thus locating where among what makes the physical truths true is that which makes the S-truths true.

Much of the literature on the debate between the *a priori* and *a posteriori* physicalists has focused, as in Witmer's paper here, on disputes about what is *a priori* knowable. In "On Ensuring that Physicalism is not a Dual Attribute Theory in Sheep's Clothing", Frank Jackson's turns his attention to identifying a distinctively *metaphysical* issue that separates the *a priori* from the *a posteriori* physicalist.<sup>37</sup> This distinction between epistemic and metaphysical issues allows for a four-fold distinction in position between *de dicto a priori* and *a posteriori* physicalism, on the one hand, and *de re a priori* and *a posteriori* physicalism, on the other. What distinguishes

*de dicto a priori* from *de dicto a posteriori* physicalism is the acceptance and the denial respectively of the claim that the truth of certain sentences, e.g. (PE) or its non-modal counterpart or, in the case of Witmer's liberal view, (L,) is knowable *a priori*.

*De re a priori* and *a posteriori* physicalism are distinguished, in contrast, by their acceptance and rejection respectively of a certain metaphysical thesis that Jackson calls the '*a priori* kind closure for the physical<sub>1</sub>' where physical<sub>1</sub> properties are just those I have been calling 'primary, physical properties'. According to that thesis, the primary, physical facts *a priori* determine all of the facts where the A-facts *a priori* determine the B-facts if and only if the nature of the A- and B-facts are such that appreciation of the former provides all of the empirical information one needs for an appreciation of the latter. The notion of *a priori* determination is usefully illustrated by one of Jackson's central examples. Consider a set of line segments arrayed on a plane. The facts about the number, orientation, and proximity to one another of those line segments *a priori* determine, in Jackson's sense, the number and kind of enclosed figures on that plane. An appreciation of the former facts provides all of the information one needs for an appreciation of the latter.

According to Jackson's *de re a priori* physicalist, the relation of the physical facts to the rest of the facts (or at least the empirical ones) is exactly like the relation of the facts about the orientation, number, and proximity of the line segments to the facts about the number and kind of figures on a plane. Here, Jackson argues that the *a priori* kind closure thesis distinctive of *de re a priori* physicalism is a required component of any metaphysical view that is to be distinctively physicalist, as opposed to a view compatible with dual attribute theories. Jackson's idea behind this claim is that while it is difficult to prove that *a priori* kind closure must be true if physicalism is, our implicit acceptance of the thesis is evidenced by our practice of looking for patterns in nature. It's because we assume that the information conveyed



by bee dances, for example, is nothing over and above the elements that make up such dances that we look to no more than the patterns found in such elements for the information that bee dances convey. But this is just to say that our practice reflects the assumption that instantiations of higher-order properties just are the instantiations of patterns we can discover by looking at the nature of the array of lower-order property instantiations. And that is just what *a priori* kind closure says about the relationship between the higher- and lower-order properties. Jackson's idea is that to suppose we need to look outside of the patterns found among the instantiations of the lower-order properties to find what determines the instantiation of higher-order properties is just to suppose that something other than the former is needed to determine the latter. Embracing that just is embracing a form of dualism. If Jackson is right, then *de dicto a posteriori* physicalists must embrace *de re a priori* physicalism if they are to avoid their view's compatibility with dualism.

Might Jackson's argument for *a priori* determination be leveraged into a justification for the physical information assumption that he seems to accept elsewhere and that Witmer here rejects? If so, then his argument for *de re a priori* physicalism allows for a justification for strict *de dicto a priori* physicalism. Here is one way it might. Suppose first that Jackson is right that to avoid compatibility with dual attribute theories, the physicalist is committed to the *a priori* kind closure of the primarily physical. Then, according to the physicalist, the nature of the facts that make 'PT' true in the liberal *de dicto a priori* thesis (L) must be such that their appreciation alone in principle allows one to know that 'E' is true, since, if physicalism is also true, what makes 'E' true will be in principle *a priori* recoverable from an appreciation of the PT-facts alone. But if that is true, then one can in principle move *a priori* from the truth of the liberal thesis (L) to the truth of the strict thesis (S).

Related to these debates between *a priori* and *a posteriori* physicalists is a debate about what is required for physical-

ism's vindication touched upon in Witmer's discussion of the location of phenomenal states among the physical ones. *De dicto a priori* and *de dicto a posteriori* physicalists both accept that defending physicalism requires showing how the non-basic truths are nothing over and above the truths about the distribution of physical properties and kinds in the primary sense (i.e. what are the basic truths, according to physicalism) by locating where among what makes the basic truths true is what makes the non-basic truths true.<sup>38</sup> *De dicto a priori* and *de dicto a posteriori* physicalists disagree, however, about what is required to solve such location problems. The *a priori* physicalists accept and the *a posteriori* physicalists reject the claim that solving such location problems requires *a priori* entailments of the non-basic truths from the basic truths plus conceptual analyses.<sup>39</sup>

The second part of Horgan's paper here addresses this issue of what is required for physicalism's defense. In particular, Horgan is concerned with whether solving physicalism's location problems requires the aid of information about the *a priori* knowable contents of the concepts associated with our non-basic terms or, more generally, the aid of semantic intuitions about under what conditions our non-basic terms apply. Horgan's view is positioned in some respects between that of the *de dicto a posteriori* physicalists, who reject the claim that vindicating solutions to physicalism's location problems requires the aid of information about the semantics of our non-basic terms and the *de dicto a priori* physicalists, who hold that not only is such information required, the required information must be *a priori* available.

According to Horgan, semantic information about our non-basic terms is required to solve location problems, but that information need not be *a priori* available.<sup>40</sup> His proposal here relies upon a semantic proposal he has developed and defended at length elsewhere.<sup>41</sup> On that semantic proposal, the semantic value of a declarative sentence S consists in its truth-conditions, where truth-conditions are cashed out in the familiar, possible worlds way, i.e. as the set of

worlds at which S is true. The novelty of his proposal consists in his holding that “very large portions of human...discourse are normally governed by semantic standards under which truth is indirect correspondence”.<sup>42</sup> In the case of a sentence S with indirect truth-conditions, it may both be that the actual world is among the set of worlds that make up S’s truth-conditions (and so that S is true) and that “there is nothing in [any of] those worlds answering to certain of [S’s] singular, predicative, or quantificational constituents”.<sup>43</sup>

What upshot does the acceptance of Horgan’s semantic proposal have for the debate over what is required to establish physicalism’s truth? Suppose, for example, that S is the sentence:

(S) ‘The State of Arizona has three public universities.’

Horgan’s idea is that the compatibility of S’s truth with physicalism’s does not require the existence of entities that answer to either ‘the State of Arizona’ or ‘...has a public university’. What their compatibility does require is that in every world that is a minimal physical duplicate of the actual world is a world at which is S semantically correct.

Although Horgan’s proposal differs from that of the *a priori* physicalists, such as Witmer and Jackson, in holding that the semantic information needed, on his view, to resolve location problems may be *a posteriori* available, it does not differ from *a priori* physicalists about the need for *a priori* semantic information in the case of locating what makes the phenomenal truths true among what makes the physical ones true. As Horgan concedes, his ontological-austerity strategy will only be of use in locating truths for which it is plausible to say that there are no entities or properties picked out by the terms in which those truths are stated. But, as Horgan says, “nothing is more indubitably real than one’s own phenomenal consciousness”.<sup>44</sup> So the ontological-austerity strategy will not provide a rival to the conceptual analysis strategy for locating phenomenal properties that Witmer pursues here and that Jackson and Chalmers have each defended elsewhere.<sup>45</sup>

## NOTES

<sup>1</sup> Thanks to Gene Witmer for helpful comments on a draft of this introduction. And many thanks especially to the Bowling Green State University Department of Philosophy and to BGSU's Social Philosophy and Policy Center for providing the funds for the BGSU conference on physicalism held in April 2005. These funds made the conference possible, and the conference, this volume.

<sup>2</sup> For more on this way of outlining the issues, see Stoljar (2001).

<sup>3</sup> "A truth" here is just a true sentence, perhaps in an idealized language.

<sup>4</sup> See Dowell (this volume).

<sup>5</sup> See for example, Hellman (1985), Crook and Gillett (2001), Melnyk (1997), Hempel (1969, 1980).

<sup>6</sup> See Dowell (this volume).

<sup>7</sup> See for advocates of this objection, see Crook and Gillet (2001).

<sup>8</sup> For ways around both of these objections to ideal physics-based formulations, see Dowell (this volume).

<sup>9</sup> See her "On Characterizing the Physical", Sec. 1.2.2, this volume.

<sup>10</sup> *Ibid.*

<sup>11</sup> See Wilson, this volume, Sec. 1.2.3.

<sup>12</sup> *Ibid.*, Sec. 2.4.

<sup>13</sup> See Dowell, objections and replies 2–5, this volume.

<sup>14</sup> See, for example, Chalmers (2004), Chalmers and Jackson (2001), and Stalnaker (2003).

<sup>15</sup> For Wilson's defense of the claim that the NFM constraint is *a posteriori*, see her 2.3.

<sup>16</sup> Dowell (this volume).

<sup>17</sup> For Worley's characterization of Montero's position, see her paper, this volume. See also Montero (1999, 2001).

<sup>18</sup> See Worley (this volume).

<sup>19</sup> See, for example, Davidson (1970).

<sup>20</sup> See, for example, Smart (1959) and, more recently, Hill (1991).

<sup>21</sup> See, for example, Jackson's (1998).

<sup>22</sup> See, for example, Boyd (1980).

<sup>23</sup> Andrew Melnyk (this volume).

<sup>24</sup> *Ibid.*

<sup>25</sup> Melnyk (2003).

<sup>26</sup> LePore and Loewer (1989).

<sup>27</sup> See Jackson (1998, p. 12).

<sup>28</sup> See Gene Witmer (this volume).

<sup>29</sup> Jackson here thanks Richard Holton for pressing this point on him. See Jackson, this volume. Jessica Wilson also defends such a view. See her (2005).

<sup>30</sup> See Witmer's "How to Be a (Sort Of) A Priori Physicalist", this volume.

- <sup>31</sup> For Jackson's argument that MPD entails PE so interpreted, see his (1998), chapter one.
- <sup>32</sup> For an example of the former, see Jackson (1998) and Jackson (2005) and for an example of the latter, see McLaughlin (2005).
- <sup>33</sup> The non-modal version is "(PMC) If PT, then A". See Witmer (this volume).
- <sup>34</sup> See Witmer (this volume).
- <sup>35</sup> This skips a couple of steps in the argument. For the full argument, see Witmer, Sec. 1.3.
- <sup>36</sup> *Ibid*, Sec. 1.4.
- <sup>37</sup> See his paper, this volume. See also his (2005).
- <sup>38</sup> See, for example, Jackson (1998) and McLaughlin (2005).
- <sup>39</sup> This is a rough statement of the *a priori* physicalist's position. For a more careful statement, see Chalmers and Jackson (2001).
- <sup>40</sup> Dowell (manuscript) also defends the claim that a posteriori semantic information may help solve location problems. There are important differences in detail, however, between the view defended there and Horgan's.
- <sup>41</sup> See, for example, his (1986a, 1986b, 1991, 1993, 2001).
- <sup>42</sup> See Horgan (this volume).
- <sup>43</sup> *Ibid*.
- <sup>44</sup> Horgan (this volume).
- <sup>45</sup> See, for example, Chalmers and Jackson (2001) and Jackson (1998). For two very different strategies for solving location problems, see McLaughlin (2001) and Dowell (manuscript). For the paper to which McLaughlin's is a response, see Horgan and Tienson (2001).

## REFERENCES

- Boyd, R. (1980): 'Materialism Without Reduction: What Physicalism Does Not Entail', in B. Ned (ed.), *Readings in Philosophy and Psychology, Volume 1* (pp. 268–305), London: Methuen.
- Chalmers, D. (2004): 'Epistemic Two-Dimensional Semantics', *Philosophical Studies* 118, 153–226.
- Chalmers, D and Jackson, F. (2001): 'Conceptual Analysis and Reductive Explanation', *Philosophical Review* 110(3), 315–360.
- Crook, S. and Gillet, C. (2001): 'Why Physics Alone Cannot Define the "Physical": Materialism, Metaphysics, and the Formulation of Physicalism', *Canadian Journal of Philosophy* 31, 333–360.
- Davidson, D. (1970): 'Mental Events', in L. Foster and J.W. Swanson (eds.), *Experience and Theory*, Amherst: University of Massachusetts Press.

- Dowell, J. (this volume): "The Physical: Empirical, not Metaphysical", *Philosophical Studies*.
- Dowell, J. (manuscript): "Serious Metaphysics and the Vindication of Reductions." <http://www.bgsu.edu/departments/phil/faculty/dowell/dowell.html>.
- Hellman, G. (1985): 'Determination and Logical Truth', *Journal of Philosophy* 82, 607–616.
- Hempel, C. (1969): 'Reduction: Ontological and Linguistic Facets', in S. Morgenbesser, P. Suppes and M. White (eds.), *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, New York: St. Martin's Press.
- Hempel, C. (1980): 'Comments on Goodman's Ways of Worldmaking', *Synthese* XLV, 193–199.
- Hill, C. (1991): *Sensations: A Defense of Type Materialism*, Cambridge: Cambridge University Press.
- Horgan, T. (1986a): 'Psychologism, Semantics, and Ontology', *Nous* 20, 21–31.
- Horgan, T. (1986b): 'Truth and Ontology', *Philosophical Papers* 15, 1–21.
- Horgan, T. (1991): 'Metaphysical Realism and Psychologistic Semantics', *Erkenntnis* 34, 297–322.
- Horgan, T. (1993): 'The Austere Ideology of Folk Psychology', *Mind and Language* 8, 282–297.
- Horgan, T. (2001): 'Contextual Semantics and Metaphysical Realism: Truth as Indirect Correspondence', in M. Lynch (ed.), *The Nature of Truth: Classic and Contemporary Perspectives*, Cambridge: MIT Press.
- Horgan, T. (this volume): 'Materialism: Matters of Definition, Defense, and Deconstruction', *Philosophical Studies*.
- Horgan, T. and Tienson, J. (2001): 'Deconstructing New Wave Materialism', in C. Gillet and B. Loewer (eds.), *Physicalism and Its Discontents*, Cambridge: Cambridge University Press.
- Jackson, F. (1998): *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford: Clarendon Press.
- Jackson, F. (2005): 'The Case for A priori Physicalism', in C. Nimtz and A. Beckermann (eds.), *Philosophy-Science-Scientific Philosophy, Main Lectures and Colloquia of GAP 5, Fifth International Congress of the Society for Analytical Philosophy, 2003* (pp. 251–265), Paderborn: Mentis.
- Jackson, F. (this volume): 'On Ensuring that Physicalism is not a Dual Attribute Theory in, Sheep's Clothing', *Philosophical Studies*.
- LePore, E. and Loewer, B. (1989): 'More on Making Mind Matter', *Philosophical Topics* 17, 175–191.
- McLaughlin, B. (2001): 'In Defense of New Wave Materialism: A Response to Horgan and Tienson', in C. Gillet and B. Loewer (eds.), *Physicalism and Its Discontents*, Cambridge: Cambridge University Press.
- McLaughlin, B. (2005): 'A Priori versus A Posteriori Physicalism', in C. Nimtz and A. Beckermann (eds.), *Philosophy-Science-Scientific*

- Philosophy, Main Lectures and Colloquia of GAP 5, Fifth International Congress of the Society for Analytical Philosophy, 2003* (pp. 267–285), Paderborn: Mentis.
- Melnyk, A. (1997): 'How to Keep the "Physical" in Physicalism', *Journal of Philosophy* 94, 622–637.
- Melnyk, A. (2003): *A Physicalist Manifesto: Thoroughly Modern Materialism*, New York: Cambridge University Press.
- Melnyk, M. (this volume): 'Realization and the Formulation of Physicalism', *Philosophical Studies*.
- Montero, B. (1999): 'The Body Problem', *Nous* 33, 183–200.
- Montero, B. (2001): 'Post-Physicalism', *Journal of Consciousness Studies* 8, 61–80.
- Smart, J.J.C. (1959): 'Sensations and Brain Processes', *Philosophical Review* 68, 141–156.
- Stalnaker, R. (2003): 'On Considering a World as Actual', in *Ways a World Might Be*, Oxford: Clarendon Press.
- Stoljar, D. (2001): 'Physicalism', *Stanford On-line Encyclopedia of Philosophy*.
- Wilson, J. (2005): 'Supervenience-based Formulations of Physicalism', *Nous* 29, 426–459.
- Wilson, J. (this volume): 'On Characterizing the Physical', *Philosophical Studies*.
- Witmer, G. (this volume): 'How to Be a (sort of) A priori Physicalist', *Philosophical Studies*.
- Worley, S. (this volume): 'Physicalism and the Via Negativa', *Philosophical Studies*.

*Department of Philosophy*  
*Bowling Green State University*  
*Bowling Green, OH 43403-0216*  
*USA*  
*E-mail: janicedowell@hotmail.com*