Physicalism and Its Prospects

by

Christian Stevens

A Thesis

Presented to

The University of Guelph

In partial fulfillment of requirements

for the degree of

Master of Arts

in

Philosophy

Guelph, Ontario, Canada

© Christian Stevens, January, 2013
ABSTRACT

Physicalism and Its Prospects

Christian Stevens
University of Guelph, 2013

Advisor: Andrew Bailey

In this thesis I explore and defend physicalism—the view that there is ‘nothing over and above’ the physical. Part of the challenge for physicalists is to make this slogan precise. They should provide a plausible account of the relation that everything must stand in to the physical in order for nothing to be ‘over and above’ it, as well as a reasonable characterization of ‘the physical’ itself. I elaborate and defend a common physicalist understanding of the ‘nothing over and above’ relation in terms of ‘global metaphysical supervenience,’ and introduce a novel strategy for characterizing the physical that sidesteps the most powerful objection to a ‘future physics’ definition of the physical—what Jessica Wilson (2006) has dubbed ‘the inappropriate extension worry.’ I then explore and respond to David Chalmers’ (1996) ‘zombie argument’ against physicalism, and Ned Block’s (2007) ‘overflow’ argument against the physicalist view of consciousness which I favour.
Acknowledgements

I would like to thank Ned Block for feedback on a paper I presented at the 2012 meeting of ‘Toward a Science of Consciousness,’ in Tucson, Arizona, which helped bring into sharp relief some of my ideas about phenomenal consciousness and cognitive access that ended up forming the bulk of the final chapter of this work. In Tucson I also met philosophers David Chalmers and Philip Goff and would like to thank them as well for their valuable input on my thoughts about the zombie argument.

My greatest debt is to my supervisor, Andrew Bailey. Professor Bailey’s insightful comments and lively encouragement at every stage of this project have been invaluable. I also owe much to my second reader, Mark McCullagh, for his careful readings of my work at important stages of this endeavor.

Finally, I am very grateful for the funding support I received from the Department of Philosophy at the University of Guelph and from the Social Sciences and Humanities Research Council.
# Table of Contents

**Introduction**

**Chapter 1: Getting Physicalism off the Ground**
- Section 1: *Formulating Physicalism*  
- Section 2: *Understanding the ‘Physical’*

**Chapter 2: Physicalism and Consciousness**
- Section 1: *The Problem of Experience*  
- Section 2: *The Zombie Argument*

**Chapter 3: Consciousness and Cognitive Accessibility**
- Section 1: *The Science of Consciousness and Overflow*  
- Section 2: *Taking Stock*

**Reference List**
Introduction

In this manuscript I provide a by no means comprehensive and tentative defense of physicalism. Physicalism is the ontological thesis that ‘everything is physical.’ It is, therefore, a form of ontological monism as opposed to dualism or pluralism. Though most of the debate surrounding physicalism has been in the philosophy of mind, issues such as understanding the ‘physical,’ and how to formulate the thesis of physicalism, have attracted a great deal of attention as well. In the first chapter I provide a defense of the most common way of formulating physicalism against some stock objections and I also provide a novel way of characterizing the physical.

In the second chapter I turn to the hotly debated question of whether or not experience is physical. I examine the famous ‘zombie argument’¹ which purports to show that experience is non-physical, and offer up two replies to it on behalf of the physicalist. I develop the second in more detail, but the first is also worth considering. The zombie argument requires a controversial link between conceivability and metaphysical possibility in order to refute physicalism, but I do not question this link. Since the question of when conceivability is a guide to metaphysical possibility (if it is at all) is one that is extremely controversial, I hold it to be a virtue of a physicalist response to the zombie argument that it does not reject this link in order to defend physicalism. Conceivability arguments play a central role in the debate over the relationship between experience and the physical. Implicit in my discussion of the zombie argument is that perhaps the importance of what may or may not seem to be conceivable has been somewhat exaggerated—though it is difficult to say exactly why or how this might be so.

If zombies are ultimately inconceivable, then experience is arguably both conceptually and ontologically reducible to the non-experiential.² I am sanguine about a conception of experience along the lines of ‘cognitive accessibility’ that seems to be implicit in the scientific method(s) for studying experience and is susceptible to an ontological reduction to the physical. In the third and final chapter, I discuss a challenge to this conception of experience and find it wanting.

² Physicalists are committed to the ontological reducibility of experience, but it is less clear that they are also committed to the conceptual reducibility of experience. This, of course, is where the thorny issue concerning the link between conceivability and possibility kicks-in. For a seminal discussion of physicalism and conceptual reducibility, see Chalmers and Jackson, 2001.
Chapter One: Getting Physicalism off the Ground

Section One: Formulating Physicalism

Physicalism is the view that ‘everything is physical,’ or, as it is sometimes also put, that ‘there is nothing over and above the physical.’ The latter rough and ready characterization of physicalism is perhaps closer to how physicalism is typically understood. The reason for this is that physicalists generally don’t hold that every property at our world is identical to some physical property, but rather, that every property supervenes upon, or is entailed by, the physical. If all the properties at our world metaphysically supervene on the physical properties, then—to borrow a metaphor from Saul Kripke (1980)—once God ‘fixes’ the physical properties she is all done; the rest comes ‘automatically.’ This is the usual sense of the physicalist’s claim that the non-physical properties of our world are nothing over and above the physical.

---

3 I formulate and typically discuss physicalism as a thesis about properties, though this is not mandatory (we could instead formulate physicalism as a thesis about ‘states of affairs,’ ‘facts,’ or ‘truths’). Of course, physicalists also believe that all events, states, and particulars are physical, or supervene upon the physical as well.

4 Actually, I cannot think of anyone who has held such a view in print. For similar ruminations on this point see Melnyk, 1996, footnote 3.

5 If $A$ properties supervene upon $B$ properties, then there cannot be a difference in $A$ properties without a difference in $B$ properties. The modal force of the ‘cannot,’ in “there cannot be a difference in $A$ properties...” ranges from nomological to logical necessity. N.B. Mere nomological necessity does not ensure that, at another possible world, $A$ properties supervene upon $B$ properties (unless we assume, controversially, that certain properties play their ‘nomic roles’ essentially. See e.g., Shoemaker, 1980; Ellis, 2001).

6 The relevant sense of entailment in play is that of metaphysical necessitation: if $B$ entails $A$, then there is no possible world in which $B$ is instantiated but $A$ is not. See Kripke, 1980. It is somewhat controversial whether or not metaphysical supervenience is sufficient for entailment. For a quick overview of some of the issues here see Bennett and McLaughlin, 2011, section 3.2.

7 It is contested whether or not the metaphysical supervenience of all non-physical properties at a world on the physical properties is enough to secure that the non-physical is nothing over and above the physical. It may be that a relation of type identity between non-physical properties and physical properties is the only way short of eliminativism in which the non-physical (ontologically speaking anyway) can be nothing over and above the physical. If this is the case, then ‘everything is physical’ and, ‘there is nothing over above the physical,’ obviously coincide. The ‘nothing over and above’ characterization of physicalism, however, at least has the virtue of leaving the matter open.
If all the properties in our world metaphysically supervene on the physical, then any other possible world that is physically indiscernible from our world will have our world’s non-physical properties as well. This suggests a natural way of cashing out physicalism:

1) Physicalism is true at a possible world \( w \) if and only if any world that is a physical duplicate of \( w \) is a duplicate of \( w \) simpliciter. ¹⁰

1), however, faces a well-known problem called the ‘epiphenomenal ectoplasm problem.’ On my own preferred way of setting out the problem, consider first a possible world \( w_1 \) where all the non-physical properties metaphysically supervene on the physical except for one, call it ‘being epiphenomenally ectoplasmic,’ that doesn’t causally interact with anything. Now, a physical duplicate of \( w_1, w_2, \) is a world where all the non-physical properties supervene upon the physical. Plausibly, all that’s preventing \( w_1 \) from being a world at which physicalism is true is the existence of epiphenomenal ectoplasm. This means that \( w_2 \) should count as a world in which physicalism is true. 1), however, rules this out; there is at least one possible world which is physical duplicate of \( w_2, \) but not a duplicate simpliciter of \( w_2, \) indeed, \( w_1, \) as we already know, is just such a world.¹¹

A common way of revising 1) to avoid the epiphenomenal ectoplasm problem is to tack on a ‘that’s-all,’ or ‘stop-right-there’ clause.¹² The idea here is that the correct ‘test,’ as it were, of the truth of physicalism at a world \( w \) lies in whether or not any other possible world that has the same physical ‘ingredients’¹³ as \( w \) and that’s all,¹⁴ is a duplicate simpliciter of \( w. \) Notice that this delivers the intuitively correct result that a physical duplicate of \( w_1 \) is a world (\( w_2 \)) at which physicalism is plausibly true; the epiphenomenal ectoplasm world is not a world that is a

---

¹¹ For the standard presentation of the epiphenomenal ectoplasm problem see Stoljar, 2009, sect. 4.3.
¹² This suggestion is due to Jackson, 1994, 1998.
¹³ I find the culinary metaphors sometimes employed in this context to be rather illuminating. See Jackson, 1994.
¹⁴ ‘That’s all’ here means nothing in addition to the physical properties and those that they entail. See Stoljar, 2010, p.135 for a brief exposition of this point.
physical duplicate of \(w_2\) and that’s all. So \(w_1\) no longer jeopardizes the claim that \(w_2\) is a world where physicalism is true. This amended version of 1) can be stated as follows:

2) Physicalism is true at a possible world \(w\) if and only if any world that is a minimal physical duplicate of \(w\) is a duplicate of \(w\) simpliciter.

A further problem, however, is waiting in the wings for 2).

Consider a possible world \(w_3\), where the relationship between the non-physical properties and the physical properties upon which they supervene is slightly weaker than that of entailment or metaphysical supervenience. Though intimate, the relationship is such that a certain kind of immaterial intervener—call it a ‘blocker’—could, were it to exist at \(w_3\), prevent the non-physical properties at \(w_3\) from being instantiated by the instantiation of the physical properties at \(w_3\). We can imagine that, at \(w_3\), the instantiation of the non-physical by the physical could only be prevented were such a blocker to exist. Now, most physicalists do not wish to count \(w_3\) as a world in which physicalism is true; giving up on the metaphysical necessitation of the non-physical by the physical is generally thought to result in some form of dualism. But, unfortunately, \(w_3\) is a world where 2) is true: any minimal physical duplicate of \(w_3\) will not be a world in which there are blockers, and so will be a duplicate simpliciter of \(w_3\).

Since the blockers problem appears to pose a serious difficulty for 2), it is worth examining whether or not an important alternative for dealing with the epiphenomenal ectoplasm problem—one put forward by David Chalmers (1996)—has the resources to rule out \(w_3\) from counting as a world at which physicalism is true. Chalmers (1996) makes the case that supervenience theses in general ought to be restricted to ‘positive’ properties, where a positive property is understood by him to be one that “…if instantiated in a world \(W\), is also instantiated by the corresponding individual in all worlds that contain \(W\) as a proper part.” This suggests the following supervenience-based formulation of physicalism:

3) Physicalism is true at a possible world \(w\) if and only if any world that is a physical duplicate of \(w\) is a positive duplicate of \(w\).\(^{18}\)

---

\(^{15}\) See Hawthorne, 2002. For this reason, the problem that follows for 2) is sometimes referred to as the ‘blockers problem.’

\(^{16}\) But for a view that takes \(w_3\) to be a world at which physicalism is true, see Leuenberger, 2008.

\(^{17}\) Chalmers, 1996, p.40.

\(^{18}\) Stoljar, 2009. Sect.4.3.
By a ‘positive duplicate’ of \(w\), what is meant is a possible world that has all the positive properties of \(w\). Now, 3) gives us the correct result that \(w_2\) is a world where physicalism is true, and \(w_1\) one where physicalism is false; a physical duplicate of \(w_1\) will be missing one of the positive properties in \(w_1\), namely, epiphenomenal ectoplasm.

But can 3) also handle the blockers problem? It has seemed to some that it can, or at least has a better chance than 2) of doing so.\(^{19}\) The reason that normally gets cited in this regard is that at least some of the positive non-physical properties in \(w_3\) will not exist in those worlds where there are blockers, and so not every possible world that’s a physical duplicate of \(w_3\) will be a positive duplicate of \(w_3\). The problem with this thought, however, is that those non-physical properties at \(w_3\) that are blocked at another possible world would seem, for that very reason, not to be positive properties in the first place.\(^{20}\) Recall that Chalmers’ (1996) definition of a positive property has it that such a property in a world \(W\) must be instantiated in all worlds that contain \(W\) as a proper part. Worlds in which a blocker exists contain \(w_3\) as a proper part, but some non-physical property instantiated in \(w_3\) won’t be instantiated in those worlds, so it’s at best unclear that the non-physical properties in \(w_3\) can be counted as positive properties. But if this is correct, then 3) arguably won’t deliver the desired result that there is some world which is a physical duplicate of \(w_3\), but not a positive duplicate of \(w_3\). The blockers problem, then, cannot clearly be averted by adopting 3) as opposed to 2).

Though the blockers problem is not easily answered, it has so far not attracted that much attention from physicalists. Part of the reason for this general lack of concern likely stems from the fact, as Daniel Stoljar (2010) points out, that issues involving negative properties, generality, and totality properties—which lie at the heart of both the epiphenomenal ectoplasm problem and blockers—are notoriously difficult to handle. This suggests that the above problems are not problems that afflict physicalism in particular.\(^{21}\) Since I am unable here to provide any kind of satisfying solution to the blockers problem, I will set it aside in favor of turning our attention to a further challenge that both 2) and 3) must also face.

It is highly plausible that either 2) or 3) captures a core necessary condition of physicalism. Whether either formulation comes close to a sufficient condition, however, is

---

\(^{19}\) See, e.g., Hawthorne, 2002, p.107.


\(^{21}\) See Stoljar, 2010, p.139.
controversial. Another important challenge to sufficiency in this context is what is sometimes called the ‘necessary beings problem.’

Consider a non-physical being that exists in all possible worlds (no doubt some theists believe God to be such a being). If such a being exists, it is natural to believe that physicalism is false. According to either 2) or 3), however, such a being is compatible with the truth of physicalism at a world \( w \) since, in an admittedly trivial way, the physical properties in \( w \) entail its existence.

An obvious way of responding to the necessary beings problem is to assert ‘Hume’s dictum.’ Hume’s dictum says that ‘there are no necessary connections between distinct existences.’ If true, it follows directly from it that there are no necessary non-physical beings; such beings would be distinct from, yet necessitated by, the physical properties of \( w \). But why believe Hume’s dictum? Well, as a thesis about ‘modal distinctness,’ it may be analytic. This interpretation of the dictum, however, will be of no use to the physicalist here.

The reason for this is that the proponent of the necessary beings problem does not, of course, think that the physical properties of \( w \) and a necessary non-physical being and its properties are modally distinct. The sort of ‘distinctness’ that the proponent of the necessary beings problem invokes is what we might call ‘weak modal distinctness,’ where a property \( G \) is weakly modally distinct from a property \( F \) if and only if there is at least one possible world in which \( G \) is instantiated and \( F \) is not, and there is no possible world in which \( F \) is instantiated and \( G \) is not. Hume’s dictum, understood as a thesis about modal distinctness (or what we might now call ‘strong modal distinctness’), is silent on the possibility or coherence of a necessary being that’s weakly modally distinct from the physical properties of a world \( w \).

A stronger gloss on the ‘distinct’ in Hume’s dictum than ‘modally distinct,’ therefore, is needed. Unfortunately, as Jessica Wilson (2010) persuasively argues, if taken to rule out weak

---

22 See Jackson, 1998.
23 See Lewis, 1986; Armstrong, 1997. Naturally, I am interested here only in the idea discussed by contemporary metaphysicians such as David Lewis and David Armstrong that they attribute to Hume, not in the historical Hume and his views.
24 Hume’s dictum is normally restricted to intrinsically typed existences.
25 A property \( F \) is modally distinct from a property \( G \) if and only if \( F \) can be instantiated and \( G \) is not, and \( G \) can be instantiated and \( F \) is not.
26 See e.g., Stoljar, 2008, p.266. For a denial, see Wilson, 2010.
modal distinctness, then Hume’s dictum becomes implausibly strong: it rules out entailment.\textsuperscript{28}

The upshot is that interpreting Hume’s dictum in a way that rules out weak modal distinctness is unmotivated. I conclude from this and the foregoing that the ‘Hume’s dictum reply’ to the necessary beings problem is a blind alley.

One might be tempted at this point to claim that if a necessary non-physical being is only weakly distinct from the physical properties in \( w \) then, contrary to our initial supposition, physicalism is not falsified by the existence of a necessary non-physical being. Most physicalists, after all, think that properties which are weakly distinct from the physical are not—in virtue of this—over and above the physical.\textsuperscript{29} In what follows, I suggest that, indeed, it is not the weak modal distinctness of a necessary non-physical being from the physical that’s problematic for physicalism, but rather how a necessary non-physical being is weakly modally distinct from the physical that causes the problem.

The way in which a non-physical necessary being is weakly distinct from the physical is, as briefly noted earlier, trivial. Since a necessarily existing being exists in every possible world, it follows that in any world in which the physical properties of \( w \) are instantiated, the necessary non-physical being is as well. So we have the physical properties in \( w \) ‘entailing’ the necessary non-physical being and its properties, but the covariation relation here is certainly not explicable in terms of determination. The physicalist, therefore, should admit that the existence of a necessary non-physical being is incompatible with the truth of physicalism.

Fortunately for the physicalist, however, there are certain considerations of a modal nature that, I will presently argue, put 2) and 3) back on track towards stating a sufficient condition for physicalism. Content to keep the discussion thus far at a relatively intuitive and technically undemanding level, I have been somewhat incautious in my presentation of the necessary beings problem; there is a glaring technical issue raised by the possibility of a necessarily existing being that, until now, I have skated over.

\textsuperscript{28} See Wilson, 2010, pp. 601-603.

\textsuperscript{29} That, at least, is what’s supposed in formulating physicalism in terms of metaphysical supervenience. Everything need not be identical to the physical in order for physicalism to be true. Being ‘realized’ by the physical in the actual world is enough. So the metaphysical possibility of, say, pain in creatures that are made entirely of silicon is consistent with the truth of physicalism at the actual world.
The issue is that, according to S5 modal logic, the possibility of a necessarily existing being implies actual necessity and, therefore, actual truth.\textsuperscript{30} The contrast, then, between the possibility of a necessary being and the actual existence of such a being flies out the window. With it also does the notion that the possibility of a necessarily existing non-physical being threatens only 2) and 3) qua putative statements of sufficiency for physicalism, and not physicalism per se.\textsuperscript{31}

We had assumed that it was both possible that there is a necessarily existing non-physical being, and possible that no such being exists. The trouble with this, however, is that—given S5—we get a contradiction: a non-physical being exists in every possible world and does not exist in a possible world, at the same time. S5, then, combined with the inherent plausibility of a possible world where a necessarily existing non-physical being does not exist, threatens to rule out the possibility of a necessarily existing non-physical being.

One drastic response on behalf of the proponent of the necessary beings problem is to reject S5. But since this reply engages modal issues of a technical nature that fall outside the scope of our discussion, I will set it aside.\textsuperscript{32} Another line of response involves first, the claim that a necessarily existing non-physical being is coherent, and second, that its coherence entails its metaphysical possibility. Some physicalists will deny the second claim, that the conceivability of a necessary non-physical being entails its metaphysical possibility, and since 2) and 3) quantify over metaphysically, not conceptually possible worlds, the mere coherence of a necessary non-physical being is of no threat to either of them. Other physicalists, meanwhile, will prefer to deny the first claim, namely, that a necessarily existing God or non-physical being is conceivable.

Though I think that both of the above options are quite defensible, for reasons that needn’t detain us here I favor the inconceivability option over the option of granting the conceivability, but denying the metaphysical possibility, of a necessarily existing non-physical being.\textsuperscript{33} To be sure, there is more that needs to be said on the matter, but, overall, the proponent

\textsuperscript{30} To give an explanation of this in intuitive terms, the only way in which it can be true at a possible world \(w\) that there is a non-physical being (say, God) in every possible world, is for it to really be true at every possible world that there is God.

\textsuperscript{31} As I had put things before, physicalism would only be falsified in this context if a necessary non-physical being were to actually exist.

\textsuperscript{32} It is also not a reply that’s obviously well motivated.

\textsuperscript{33} For a defense of the inconceivability of a necessarily existing God in a different context, see Chalmers, 2010, pp.177–179.
of the necessary beings problem seems to be in an unenviable position: as we know, a possible world in which there is a necessarily existing non-physical being is logically incompatible with a possible world in which there is no such being, and the latter is a good deal more antecedently plausible than the former.\(^{34}\)

It may be argued, however, that the necessary beings problem is just one example of a possible world \(w\) where everything is entailed by the physical, and there are properties in \(w\) that are something over and above the physical (i.e., physicalism is false) even though determined by the physical. Take for instance a possible world \(w_4\), in which 2) or 3) is true, but some mental properties in \(w_4\) are, nevertheless, something over and above the physical properties in \(w_4\). Let’s label the view that \(w_4\) is coherent or possible ‘necessitation dualism.’\(^{35}\)

While I certainly cannot discern any obvious contradiction in \(w_4\) as stated, I do have a difficult time imagining it. I always seem to end up imagining a world where all the properties are determined by the physical, not a world where some mental properties are ‘emergent,’\(^{36}\) or metaphysically distinct from the physical in a manner that suffices for them to be something over and above the physical. Karen Bennett (2008) agrees with this sentiment, writing “…I find it quite hard to make sense of a version of dualism according to which the mental is both genuinely distinct from and metaphysically necessitated by the physical.”\(^{37}\) This isn’t yet an argument against the coherence or possibility of \(w_4\), but, as Bennett (2008) points out, proponents of necessitation dualism such as Frank Jackson (2006) and Andrew Melnyk (2003) take its truth to be self-evident, and this deserves to be called into question. So far then, our discussion of necessitation dualism has been inconclusive. Before turning our attention elsewhere, however, there is another issue in the vicinity that bears mentioning.

A common criticism of supervenience-based formulations of physicalism is that supervenience is just a relation of property covariation, and if a set of \(A\) properties covaries with a set of \(B\) properties, then the metaphysical relation between \(A\) properties and \(B\) properties

\(^{34}\) It is, for one thing, far easier to imagine the latter than the former. See Chalmers, 2010, p.178.
\(^{35}\) See Stoljar, 2010, Ch.8.
\(^{36}\) In my view, emergence only obtains when there is nomological, not metaphysical supervenience (see Van Cleve, 1990).
\(^{37}\) Bennett, 2008, p.299.
remains to be explained. Therefore, supervenience-based formulations of physicalism such as 2) and 3) do not state a sufficient condition for physicalism. \(^{38}\)

In response, I think that the physicalist can accept that supervenience is just a relation of property covariation, and that it is not a brute fact that two sets of properties covary with each other, while, nevertheless, denying that supervenience-based formulations of physicalism are threatened by this. Take, for instance, Melnyk’s (2003) view that the metaphysical supervenience of the non-physical on the physical is best explained by the realization of the non-physical by the physical. Assuming the falsity of necessitation dualism, it is plausible that metaphysical supervenience guarantees realization, which, in turn, according to Melnyk (2003), would mean that physicalism is true. \(^{39}\) More generally (and once again setting aside necessitation dualism) it is plausible that whatever explanation there is for the metaphysical supervenience of the non-physical properties in a world on the physical properties, it will turn out to be one that’s acceptable to the physicalist. \(^{40}\)

I conclude that in spite of the challenges supervenience-based formulations of physicalism such as 2) and 3) face, they do enough to earn their keep.

---


\(^{39}\) Melnyk, 2003, as mentioned earlier, endorses necessitation dualism. This is why he denies that metaphysical supervenience guarantees realization.

\(^{40}\) Horgan, 1993, thinks that we need an explanation of metaphysical supervenience that is physicalistically acceptable. My claim is that metaphysical supervenience guarantees this. Again, the disagreement here largely centers on the issue of necessitation dualism. Horgan, 1993, believes necessitation dualism is true, while I do not.
Section Two: Understanding the ‘Physical’

In the previous section I tentatively defended supervenience-based formulations of physicalism such as 2) and 3). Before we can hope to evaluate whether or not ‘everything supervenes upon the physical,’ however, there is, among others, the important question of what is meant by the ‘physical’ in this context.

Although some conception of the physical is required to sensibly engage the issue of whether or not physicalism is true at our world—or any world for that matter—we need not attempt to provide necessary and sufficient conditions for being physical to enter this debate.\(^{41}\) That being the case, since we can (apparently) readily identify things that are clearly physical (e.g., rocks and trees) as well as things that are clearly non-physical (e.g., ghosts and angels), one might be tempted to suppose that this is already enough for an adequate understanding of the physical in physicalism.

Unfortunately, however, this conception of the physical is on a collision course with modern physics. As Bertrand Russell (1961) once put it, with the advent of modern physics, “…matter has become as ghostly as anything in a spiritualist séance.”\(^{42}\) Take, for example, the elementary particles of the ‘standard model.’ They are all point-like. A couple of them (photons and gluons) have no mass, and one of them (the neutrino) can apparently pass through just about anything. So something’s got to give here.\(^{43}\) But it won’t be the notion that the entities and properties of the standard model are to be counted as physical; therefore we are in need of a different conception of the physical.\(^{44}\)

A more promising suggestion, perhaps, is that an entity or property is physical if and only if it is one of the entities or properties needed to provide a complete account of prototypical examples of the physical (e.g., rocks and trees).\(^{45}\) Let us call this the ‘object-based’ conception of the physical.\(^{46}\) The object-based conception, however, does have the troubling consequence that if there exists a fundamental particle that does not figure in a complete account of prototypical examples of the physical, we are in need of a different conception of the physical.

\(^{41}\) At least I will assume as much in what follows.

\(^{42}\) Russell, 1961, p.104.

\(^{43}\) That is to say, modern physics and our ‘common-sense’ understanding of the physical are incompatible.

\(^{44}\) cf. Montero, 2009.


examples of the physical,\textsuperscript{47} then it cannot be counted as physical. Additionally, if irreducible experience\textsuperscript{48} were to figure in a complete account of prototypical examples of the physical, then it would be counted as physical. The object-based conception of the physical, then, is in one sense too restrictive, and in another, overly permissive.

More common ‘physics-based’ characterizations of the physical have similar problems. Suppose there are ‘fundamental’\textsuperscript{49} properties that are inaccessible in principle to physics. A physics-based conception of the physical—\textit{ab initio}—rules out such properties from being counted as physical. Also, a characterization of the physical in terms of a future (ideal) physics must contend with what Wilson (2006) has dubbed the ‘inappropriate extension worry:’ for all we know, a future (ideal) physics might include irreducible experience.\textsuperscript{50}

Of course, if the notion of irreducible experience is ultimately incoherent, or coherent but not metaphysically possible, then both the future (ideal) physics conception of the physical, and the object-based conception are, in spite of their shortcomings, more or less adequate conceptions of the physical. This is good news for physicalists, as there is at least one way to argue for the incoherence or impossibility of irreducible experience that does not presuppose either conception.

The way I have in mind is to argue that experience metaphysically supervenes upon the dispositional.\textsuperscript{51} I take it that this is essentially the position of ‘traditional’ physicalists. After all, it is this kind of reducibility claim that faces the ‘explanatory gap’\textsuperscript{52} and ‘hard problem of consciousness.’\textsuperscript{53} It is also, of course, the target of conceivability arguments against physicalism such as the ‘zombie’\textsuperscript{54} and ‘knowledge’\textsuperscript{55} arguments.\textsuperscript{56} One may reasonably wonder then why we

\textsuperscript{47} The axion is one such candidate. cf. Montero, 2009, p.179.
\textsuperscript{48} By ‘irreducible experience’ here I mean experience that is not metaphysically necessitated by any non-experiential property. cf. Montero, 2006.
\textsuperscript{49} An entity is ‘fundamental,’ if and only if it exists at a relatively low level of constitutional complexity— which is arguably the subject matter of physics (See Wilson, 2006). So I distinguish between fundamentality and irreducibility.
\textsuperscript{50} A ‘current physics’ characterization of the physical doesn’t have this problem, but it has many others. Most damaging perhaps is the incompleteness of current physics: the entities and properties of some future physics cannot be counted as physical, and some of them will not, in all likelihood, supervene upon the physical—rendering physicalism false.
\textsuperscript{51} Where dispositions and functional-roles are understood as rigid designators.
\textsuperscript{52} See Levine, 1983.
\textsuperscript{53} See Chalmers, 2010.
\textsuperscript{54} See Chalmers, 1996.
shouldn’t simply characterize the physical as follows: an entity is physical if and only if it figures in a future (ideal) physics, and has a dispositional essence. Let’s refer to this as the ‘power-based’ conception of the physical.

Though there are arguments in favor of the view that physics wholly characterizes its entities in terms of their dispositions or theoretical roles,\textsuperscript{57} as a fairly thoroughgoing naturalist, I prefer to avoid setting \textit{a priori} limits on the sciences where possible. Moreover, if an entity were posited by a future (ideal) physics that did not have a dispositional essence, I see no persuasive reason for thinking it non-physical by virtue of this alone. Therefore, separating out the standard physicalist view on the reducibility of experience (enshrined in the power-based conception of the physical) from the issue of characterizing the physical potentially affords a more flexible understanding of the physical.

My strategy does, however, involve a substantive commitment with respect to the nature of fundamental entities: it rules out the view, as Sungho Choi and Michael Fara (2012) put it, that disposition ascriptions are “…merely convenient ways of talking about \textit{categorical properties}.”\textsuperscript{58} But this does not imply that my account commits one to the view that fundamental entities have dispositional essences, as I allow that a future (ideal) physics may countenance categorical properties as well. So the proposal has the virtue of remaining neutral on the issue of whether or not the dispositional properties of fundamental entities require, or have, categorical bases. The metaphysical picture that is being supposed here then, is that ‘categoricalism’\textsuperscript{59} is false, and either a ‘mixed-view,’\textsuperscript{60} or ‘dispositionalism,’ is true.\textsuperscript{61}

\textsuperscript{55} See Jackson, 1982,1986.
\textsuperscript{56} It is important to remember that the conclusion of both the knowledge argument and the zombie argument is arguably a disjunctive one: either experience is irreducible, or it metaphysically supervenes upon some irreducible ‘neutral stuff’ (which, of course, is non-dispositional) that we will likely never fully comprehend—such as Chalmers’ (1996) ‘proto-experience.’ Those who wish to dispute the conclusion of the conceivability arguments, then, cannot simply substitute non-experiential for physical (pace Stoljar, 2006).
\textsuperscript{58} Choi and Farah 2012, section 3 (their italics).
\textsuperscript{59} Categoricalism is the view that fundamental properties have intrinsic essences. See e.g., Armstrong, 1997; Lewis, 2009.
\textsuperscript{60} A mixed-view holds that every fundamental property is in some sense both dispositional and categorical. This type of view is often associated with the work of C.B. Martin and John Heil. See e.g., Martin and Heil, 1999.
One problem that might be raised for my proposal at this point is that if a mixed-view is correct, then it’s unclear whether a future (ideal) physics will make reference to categorical properties. This is pressing given the widespread acceptance of the view that physics wholly characterizes its entities in terms of their dispositions. I agree that current physics characterizes its entities wholly in terms of their dispositions or (better) theoretical roles. But I disagree with the suggestion that a future (ideal) physics must or very likely will do so as well. It’s more in the spirit of physicalism, as I put it earlier, to avoid setting a priori limits on the scope of scientific theories. Granted, however, that if a future (ideal) physics characterizes its entities wholly in terms of their dispositions, and the mixed view is correct, then my proposal for characterizing the physical may well be rendered unsatisfactory. I take solace in knowing that properties that are inaccessible in principle to physics are problematic for other fairly promising attempts at characterizing the physical as well.62

In a nutshell, then, my proposal for characterizing the physical is to hold that the physical is what figures in a future (ideal) physics, while denying that irreducible experience may be counted as physical, since it metaphysically supervenues upon the dispositional properties of fundamental entities. Though this rules out categoricalism and versions of ‘physicalism’ which appeal to intrinsic properties at the fundamental level to ground experience, this strategy should appeal to those physicalists who do not wish to commit to dispositionalism, but, nevertheless, think that experience supervenues upon the dispositional properties of fundamental entities.

Though the strategy I have advanced for understanding the physical requires further elaboration and defense, I trust that it is sufficiently clear. Of course, it does stand or fall on the question of whether or not experience supervenues upon the dispositional. But since the traditional physicalist is, in any case, committed to the metaphysical supervenience of the experiential upon the dispositional, this should not trouble her. At any rate, I will now turn my attention to the question of whether or not experience supervenues upon the dispositional.

61 A complication here is that it’s by no means trivial that a mixed-view is coherent, or can avoid collapsing into one of its relevant rivals (i.e., categoricalism or dispositionalism). If a disposition and its base are not identical, then a ‘categorical base’ of a disposition may, on a certain understanding of ‘categorical,’ fail to be non-dispositional. I do not attempt to settle the matter since it raises vexed issues such as how the intrinsic vs. extrinsic and relational vs. nonrelational distinctions are best understood.

62 See e.g., Wilson, 2006.
Chapter Two: Physicalism and Consciousness

Section One: The Problem of Experience

A mental state is experiential if and only if there is ‘something it is like’ to have, or be in it. It is widely agreed that the existence of such states poses a serious challenge to the truth of physicalism. As Chalmers (2010) puts it, “[f]or any physical process we specify there will be an unanswered question: why should this process give rise to experience?” If we can wholly characterize experience in terms of its ‘causal-role,’ then there’s no great mystery as to how some ‘physical process’ might ‘give rise to experience.’ But, intuitively, experience cannot be wholly characterized in terms of its functional role. Given any functional specification of experience, we always seem able to coherently ask: why should this functional-role be experiential?

Take, for example, Ned Block’s (1995) notion of ‘access-consciousness.’ According to Block (1995), a mental state is access-conscious if and only if “…it is poised for free use in reasoning and for direct ‘rational’ control of action and speech.” Suppose that access-consciousness is experience. We can then coherently ask, why should a mental state that’s poised for free use in reasoning and rational control of action and speech be an experiential one as well? In other words, why should there be something it is like to have a mental state that’s access-conscious? Surely it is at least coherent that a creature could have a mental state that’s access-conscious without it also being experiential. We have now moved from puzzlement to argument. In the following section, I turn to a discussion of the major conceivability argument against physicalism: the zombie argument. For ease of exposition I will use the term ‘physical’ in discussing the argument. It should be remembered, however, that the power-based conception of physical that I will be operating with here is what I referred to as ‘the dispositional’ in the previous chapter.

63 See Nagel, 1974.
65 Block, 1995, p.382.
**Section Two: The Zombie Argument**

If we can always coherently ask why some physical property generates experience, or is experiential, then, strictly speaking, it is coherent that a physical duplicate of the actual world could be one in which experience is absent.\(^6\) Chalmers (1996) dubs a world that is a physical duplicate of the actual world, but where experience is absent, a ‘zombie world;’ and creatures that are physical duplicates of us but do not have conscious mental states, ‘zombies.’ Andrew Bailey (2009) points out that it is natural and helpful to think of a zombie world as what’s left over after experience is subtracted from the actual world, while leaving its physical properties unchanged.\(^7\) To assert the coherence of a zombie world is to assert the logical possibility of this subtraction. To deny its coherence is to deny that we can subtract experience while leaving the physical properties of the actual world unchanged.

If a zombie world is coherent, then not only is experience not identical to the physical, but it also does not metaphysically supervene upon the physical. We can summarize the argument as follows:

1. A zombie world is conceivable.
2. If a zombie world is conceivable, then a zombie world is metaphysically possible.
3. If zombies are metaphysically possible, then physicalism is false.
   
   Ergo,
   
   4. Physicalism is false.

If a zombie world is coherent/metaphysically possible we may wonder whether or not non-physical experience in our world can cause certain physical events to occur, or if experience is causally inert. It is tempting to suppose that if zombies are coherent, then interactionism is true at our world. The trouble with this thought, however, is that, given interactionism, a physical duplicate of the actual world in which experience is absent does not seem to be a complete physical duplicate of the actual world; those physical events caused by experience in the actual world apparently won’t obtain.\(^8\) For this reason, it is sometimes alleged that the coherence of a zombie world implies that experience is epiphenomenal. Some view this as a *reductio* of the

---

\(^6\) And vice-versa. I briefly elaborate on this point later on.  
\(^7\) I am paraphrasing here from Bailey, 2009, p.131.  
\(^8\) That is, unless we accept systematic causal overdetermination.
zombie argument: we know that experience is not epiphenomenal in the actual world, and so if a zombie world has this consequence, then, ultimately, it is incoherent.

I am unsure whether or not the coherence/possibility of a zombie world implies epiphenomenalism in the actual world. I also don’t know whether epiphenomenal experience in the actual world is as likely to be false as some argue that it is.69 I would like to suggest, however, that even if the proponent of the zombie argument—who we can refer to as a ‘zombist’—can succeed in making room for the truth of interactionism at our world, this maneuver is not a particularly helpful one.

First, there is good inductive support for what’s called the ‘causal closure of the physical,’ the hypothesis that every physical event has a sufficient physical cause.70 Second—and this is what really matters here—the empirical evidence for the causal closure of the physical strongly suggests that the truth of interactionism at our world is a question that’s empirically tractable.71 Anti-realist considerations aside, we may one day be in a position to know whether or not causal closure is true. So if the coherence of a zombie world implies either the truth of epiphenomenalism or interactionism at our world, then the physicalist is arguably in a fairly robust dialectical position here: there seem to be good arguments against experience being epiphenomenal in the actual world,72 and our completed sciences will evince whether or not causal closure is true—so we may one day be in a position to know whether or not a zombie world is coherent! The advantage of this response to the zombie argument over others, then, is that it potentially turns the coherence of a zombie world into an empirical question.73

Another line of response to the zombie argument that I find promising is sometimes called the ‘reverse-zombie argument.’74 Imagine a minimal physical duplicate of the actual world in which experience is present. If such a world is conceivable, then it is also possible. If it is

---

69 See, e.g., Bailey, 2006.
70 Causal closure can also be sensibly weakened as follows: every physical event that has a cause has a sufficient physical cause.
71 *Modulo* causal overdetermination. Overdetermination, however, leads back into a kind of epiphenomenalism, so there’s no use in appealing to it in this context. See Bailey, 2006, 2009.
72 See Bailey, 2006.
73 A more ambitious line of response to the zombie argument is to argue that if the coherence of a zombie world implies epiphenomenalism at any world, then it is ultimately incoherent (see e.g., Kirk, 2005). I will not pursue this here.
74 See Frankish, 2007; Brown, 2010.
possible, then experience is either identical to the physical, or it is entailed by the physical. Therefore, experience is nothing over and above the physical. Following Richard Brown (2010), let’s call a world that’s a minimal physical duplicate of the actual world in which experience is present, a ‘shombie world.’ And let’s call our minimal physical duplicates that have conscious mental states, ‘shombies.’ Since zombies and shombies cannot peacefully coexist, the zombist is obliged to say something about the (alleged) coherence of a shombie world.

A zombist might object that asking us to imagine a shombie world is the same thing as asking us to imagine a world in which physicalism is true. While this may be true of some reverse-zombie arguments, it is not true of mine. Conceiving of a world that is a minimal physical duplicate of the actual one in which experience is present is importantly different from conceiving of a world in which experience is metaphysically necessitated by the physical; the former involves conceiving of a qualitative non-modal state of affairs, while the latter involves conceiving of a necessity within a world. I only endorse the conceivability of a world that is a minimal physical duplicate of the actual one in which experience present. The conceivability/possibility of such a world implies that experience is necessitated by the physical in the actual world—which, of course, is exactly what the proponent of the shombie argument is after. The shombie argument, then, as I have presented it here, does not involve conceiving of the truth of physicalism at a world. It therefore avoids the problematic issues that surround conceiving of what is possible or necessary within a world. With this point in place, it’s not obvious what the zombist can say in response to the shombie argument except that zombies are conceivable on ideal rational reflection, while shombies are only prima facie conceivable. This line of response, however, seems difficult to sustain.

The apparent conceivability of a shombie world suggests to some that there is something wrong with the zombie argument—even if we can’t specify where the problem with it lies. David

---

75 There’s no other way for experience to exist in such a world (because we are postulating a minimal physical duplicate of the actual world. It’s important to note that we can also provide a ‘reverse-knowledge’ argument as well (see Brown, 2010). The discussion that follows, therefore, applies mutatis mutandis to Jackson’s, 1982,1986, knowledge argument.
76 If a shombie world is possible, then experience is either identical to, or metaphysically supervenes upon, the physical. This is parallel to the zombie argument. If zombies are possible, then this implies that consciousness is irreducible. In other words, there won’t be any world in which there are only physical properties and consciousness exists (i.e., no shombie worlds!).
Beisecker (2010), for example, thinks that reverse-zombie arguments “…resemble Gaunilo’s rejection of the ontological argument.” Though tempting, this is premature. A more charitable and plausible take on the situation is that while zombies and shombies are both prima facie conceivable, only one of them will turn out to be ideally conceivable.

A zombist might suggest that the impression some have that the zombie argument must go wrong somewhere if shombies are conceivable is due to the fact that the shombie argument stacks the deck in favor of physicalism: it assumes that experience is functionally analyzable. But while it may be true that the shombie argument implies that experience is functionally analyzable, it no more presupposes this than the zombie argument presupposes that it is not. A zombist might respond by claiming that even if the shombie argument does not beg the question, it is still worse off than the zombie argument because we have strong independent reasons for thinking that experience is not functionally analyzable. I do seem to have a kind of brute intuition that, say, life is functionally analyzable, but experience is not. Apart from that, I’m not quite sure what the zombist has in mind here. Perhaps he thinks of the explanatory gap and hard problem as a strong independent source of evidence for an absence of analysis when it comes to experience. I disagree. The suggestion that I can always coherently ask why some physical process gives rise to experience, strictly speaking, depends upon the outcome of the conceivability arguments. If zombies aren’t conceivable, then I can’t coherently ask why this physical process gives rise to experience. That would be like asking: why does this physical process give rise to water?

So, at best, there is some intuitive support for the coherence of a zombie world over the coherence of a shombie one. But this can hardly be used as a premise in an argument against the coherence of a shombie world. We seem, then, to have a stalemate here that will likely degenerate into table pounding. What we need are new thought-experiments and more science. But what kind of thought-experiment(s) could resolve or advance this debate?

Let us wind the clock back to before Edmund Gettier (1963) presented his famous counterexamples to the theory that knowledge is justified true belief. Some might have claimed that having a justified true belief is not a sufficient condition for knowledge since a justified true belief that is not knowledge cannot be ruled out a priori. They might have even gone a step

---

further and claimed that one can imagine a situation in which Fred has a justified true belief that P, but does not know that P. Incensed by the argument, defenders of the view that having a justified true belief is sufficient for knowledge might have retorted that, just as one might imagine a situation in which Fred has a justified true belief, but not knowledge, one can also imagine a situation in which Fred has a justified true belief, and knows that P. Thank goodness for Gettier (1963). If it is thought that the prospects for understanding the essence of knowledge are grim nowadays, one need only consider the epistemic situation of those imaginary interlocutors just described to see how far we’ve come. Unfortunately, the (imagined) pre-Gettier debate I have sketched seems to be more or less the situation we are in with regard to experience. So we await a conceptual revolution à la Gettier.

I have outlined two responses to the zombie argument that both suggest we are simply unable to know, from our current epistemic situation, whether or not a zombie world is conceivable. The state of play, as I see it, is one of evidential equality: the intuition(s) that consciousness is not functionally analyzable is neither trumped by nor trumps the successful track record of the sciences. Moreover, it’s not as though there aren’t any plausible functional analyses of consciousness currently on offer. Most promising, in my view, are those along the lines of Block’s (1995) notion of access-consciousness. I defend this suggestion against an important rival in the following chapter.

---

81 It is, of course, natural to view the situation as a clash between what Wilfred Sellars (1963) called the ‘manifest’ and ‘scientific’ images.
Chapter Three: Consciousness and Cognitive Accessibility

Section One: The Science of Consciousness and Overflow

Recent methods for distinguishing conscious from unconscious perception emphasize subjective reports—such as ratings of stimulus visibility. This stands in contrast to a long tradition in psychophysics (grounded in signal-detection theory) that favors the view that subjects are not conscious of a stimulus unless they can perform at above chance levels on some direct task of stimulus detection or classification. A major problem with this methodology for studying experience, however, is that it seems to overestimate conscious perception: there are many conditions in which subjects perform at above chance levels, yet deny perceiving the stimulus. But why should we consider report and performance to be evidence for experiential mental states in the first place?

The most sensible answer to this question, I think, is that we are assuming a conception of experience where experiential mental states have certain dispositions. The thought is that awareness of a stimulus x normally enables us to classify and/or detect x at a high level of accuracy, and awareness of a stimulus x also normally enables us to report having seen x. A general underlying conception of experience that unites reports and well-above-chance-level performance as evidence for it is Block’s (1995) notion of access-consciousness. Recall that, according to Block (1995), a mental state is access-conscious if and only if “…it is poised for free use in reasoning and ‘rational’ control of action and speech.” Rational control of action and speech includes reports, unprompted button pushes in response to stimuli, et cetera. In other words, the notion of access-consciousness encompasses those dispositions that are needed in

---

82 This problem also arises, though to a much lesser extent, when stimulus-detection or classification performance is combined with confidence judgments such as post-decision wagering (e.g., would you bet that you were correct?). See Kanai et al., 2010.
83 Just as objective behavioral measures of consciousness can often overestimate conscious perception, so can those that only employ subjective report. For example, if a subject reports having seen a stimulus quickly followed by a ‘mask’ on every trial of an experiment, but only half of the trials actually involved a stimulus being flashed on screen prior to the onset of a masking stimulus, then we have (considerable) reason to doubt that she was conscious of the stimulus on each trial—he or she may have simply been guessing all along. So it is best to combine subjective report and objective behavioral measures—which is standard practice in the scientific study of consciousness.
84 Block, 1995, p.382.
order for report and performance to count as evidence for the existence of conscious states, if report and performance are to count as evidence for such states at all.

This has important implications for the debate over whether or not there are experiential mental states that are not available for report. I can discern no contradiction in the notion of experiential mental states that are not access-conscious, but since report and objective performance measures are plausibly only evidence for states that are access-conscious, it is difficult to see how there could be an empirical case for experiential mental states that are not ‘cognitively accessible.’ Block (2012, in conversation), who has previously defended the view that there is empirical support for experience that is not cognitively accessible, does not try to show that report or objective performance can serve as evidence for experiential mental states that are cognitively inaccessible. Instead, he suggests that his so-called ‘overflow’ argument has always been based on first-person evidence that ‘we see more than we can report’ in such tasks as the one used in the famous experiments of George Sperling (1960). I am prepared to accept

---

85 So an apparent report about a state that is not access-conscious cannot obviously be explained in terms of that state, since it is, _ex hypothesi_, not in a position to be reported.
86 See e.g., Block, 2007.
87 Sperling (1960) presented subjects with a visual array of letters (e.g., three rows of four letters each) for a brief duration (typically 50ms). In the so-called ‘whole-report’ condition, he had subjects report as many items from the array as they could in their proper spatial locations. Under whole-report, he found that subjects could accurately report about 4 letters. Sperling (1960) noted, however, that subjects often reported having seen more than the 4 letters that they could recall. In an effort to tap into this intuition, Sperling (1960) devised the ‘partial-report’ condition in which one row of the array was cued for report after a short delay (a high-pitched tone to indicate that the top row was to be reported, a medium-pitched tone to indicate the middle row _et cetera_). What Sperling (1960) found was that subjects could, on average, report almost all of the letters from the cued row. He argued on the basis of this finding that after array offset, subjects must have had 9.1 items (76%) of the full array available to them, in contrast to the 4.5 items reported under the whole report condition. Sperling (1960) also found that partial-report remained superior to whole-report even when the partial-report cue was delayed for as long as 500ms after array offset. But after about 1 second, he found that partial-report performance dropped down to the level of whole-report. Sperling (1960) had apparently discovered a high capacity rapidly decaying memory store. Moreover, subsequent work (e.g., Averbach and Sperling, 1961) showed that this short-lived memory store was also highly sensitive to exposure parameters. For instance, when a high luminance field preceded and followed the array, partial-report performance dropped to the level of whole-report after 500 ms. But when the array was preceded and followed by a dark field Averbach and Sperling (1961) found that partial-report performance was superior for up to 5s. This suggested that the high capacity short-lived memory store in question was also ‘pre-categorical’ in nature: a kind of primitive ‘visible copy,’ or ‘snapshot’ that maintains the precise metric of the display. Subsequent work has shown,
that there is first-person evidence in favor of overflow. But I am not prepared to grant Block’s (2007) proposed ‘mesh’ between psychology and neuroscience on the basis of our phenomenological evidence that experience outstrips cognitive accessibility. I will argue that the price we pay in refusing to credit the intuition that we see more than we can report in the Sperling (1960) task (and elsewhere) seems to be roughly the same as, if not lower than, the price we must pay if we accept Block’s (2007) framework.

As mentioned earlier Block (2007) thinks that the experiments by Sperling (1960) and others using Sperling’s (1960) partial-report procedure provide evidence for experiential mental states which are not cognitively accessible. It’s important to remember, however, that Block (in conversation) has not urged that subjects’ reports and performance in these experiments are evidence for experience without cognitive accessibility. Rather, he suggests that we have strong first-person phenomenological evidence for experience without accessibility in these conditions. If we accept this, then Block (2007) suggests that we should also accept the hypothesis that a certain type of neural processing (so-called ‘recurrent’ or ‘reentrant’ processing) which correlates with both (putative) experience that is not cognitively accessible, as well as experience which is cognitively accessible, is the core neural basis of experience. Block’s (2007) framework, then, has it that the neural machinery of cognitive accessibility is distinct from the neural machinery of experience. One of the biggest intuitive drawbacks of Block’s (2007) framework, however, is that it counts phenomena such as ‘inattentional blindness’ as instances of conscious perception—since the same type of neural processing is present in such cases.\textsuperscript{88}

Inattentional blindness refers to “…the failure to notice a fully-visible, but unexpected object because attention was engaged on another task, event, or object.”\textsuperscript{89} In a famous study by Daniel Simons and Christopher Chabris (1999), participants watched a video of two groups of people passing around a basketball and were asked to count the number of passes made by people in white shirts and ignore the passes made by those wearing black. Remarkably, “[u]nder these conditions, approximately 50% of observers failed to notice when a person in a gorilla suit entered the display, stopped and faced the camera, thumped its chest, and exited on the far side

\textsuperscript{88} See Lamme, 2006, p. 498, for discussion.
\textsuperscript{89} Simons, 2007, p.1.
of the display." Now the problem with counting cases of inattentional blindness as cases of ‘inattentinal amnesia,’ is that there is no phenomenological evidence for having seen the unattended item. When I do these experiments and miss some clearly visible item due to inattention, it’s not as though I simply don’t report it because I am, say, unsure whether or not I saw something. I literally, from the inside, have no clue that I’ve missed anything until I do the experiment over again and don’t follow the instructions. So if we’re saving some intuitions about experience (under certain experimental conditions) by adopting Block’s (2007) framework, we’re sacrificing others. Moreover, there is now evidence suggesting that recurrent processing, which, remember, Block (2007) suggests is the core neural basis of experience, also occurs when subjects are under general anesthetic, in a dreamless sleep, and during epileptic absence.

Overall then, it seems to me that a functional analysis of experience involving cognitive accessibility is not undermined by phenomenological evidence for experience without cognitive accessibility. That is, unless it is the case that experience which is not cognitively accessible is (ideally) conceivable. Naturally, I deny this. I can imagine a minimal access-conscious duplicate of me that’s also a minimal phenomenal duplicate. At least, I think I can. So my response here runs parallel to the one I gave earlier to the zombie argument: we simply don’t know, from our current epistemic situation, whether or not experience that’s not cognitively accessible is coherent. The fact that there is also a robust neural account of cognitive accessibility—the so-called ‘global neuronal workspace model’ makes me even more inclined to think that it would be a serious error to reject the view that experience involves cognitive accessibility solely on the basis of a phenomenological intuition, or a phenomenological intuition and a sketchy mesh between it and neuroscience.

It will surely be complained that I have not so much made a case for physicalism and an analysis of experience involving cognitive accessibility as shown that these options are no more likely to be true or false than their relevant rivals. There is some truth to this. I said earlier that I

---

90 Simons, 2007, p.1. It should be noted, however, that while subjects often do not report having seen the unexpected stimulus, they do show some ‘priming’ from it. Mack and Rock (1998), for example, noted that immediately following an inattentional blindness experiment involving words, subjects had a tendency to complete word fragments with the unreported word as opposed to other more common words that were not used in the experiment.

91 See Wolfe, 1999.

92 See Arthuis et al., 2009.

93 See Dehaene and Naccache, 2001. For a recent review, see Dehaene and Changeux, 2011.
think the current evidence doesn’t favor the conceivability of shombies over zombies, and vice versa. This does not mean, however, that, all things considered, physicalism is no more likely to be true than dualism. The same goes for my defense of an analysis of experience involving cognitive accessibility. At present, it may be no more likely that a minimal access-conscious duplicate of me that’s also a minimal experiential duplicate is conceivable than a minimal experiential duplicate of me that’s not a minimal access-conscious duplicate is. But this does not mean that, all things considered, a functional analysis of consciousness that involves cognitive accessibility is no more likely to be true than one that analyzes consciousness in terms of some neural property, such as recurrent processing. By blunting the force of the zombie argument and Block’s (2007) overflow argument, I put physicalists who think that an analysis of experience must involve cognitive accessibility in a stronger dialectical position than before with respect to their rivals.

---

94 Remember that the coherence of a shombie world entails physicalism, while the coherence of a zombie world entails dualism, so evidence in favor of one or the other is not the same thing as evidence for physicalism or dualism.

95 This may, of course, sound strange. The idea of an analysis of experience in terms of some non-psychological property may strike some as being pitched at the wrong conceptual level. Another reason this might sound strange is that physicalists who deny that we can analyze experience in terms of ‘structure and dynamics’ (see Chalmers, 2010) tend to go on to identify experience with some neural property. Finally, ‘functionally analyzable’ suggests to many that a ‘functionalist’ (i.e., inputs-outputs and relations to other mental states) analysis is intended. By ‘functionally analyzable,’ however, I just mean analyzable in terms of structure and dynamics.
Section Two: Taking Stock

In the first chapter of this thesis I defended supervenience-based formulations of physicalism against certain well-known objections, and proposed a novel strategy for characterizing the physical. This strategy involves a conditional: if experience metaphysically supervenes upon the dispositional, then we can (satisfactorily) characterize the physical by reference to a future (ideal) physics. One of the main attractions of this strategy is that it allows categorical bases to be counted as physical, but avoids the possibility of those bases being experiential properties.

In the second chapter I briefly introduced the problem of experience, and presented two replies to the well-known zombie argument against physicalism. The second one, which I spent more time on, is the so-called ‘reverse-zombie’ argument. Having examined a number of replies to the reverse-zombie argument on behalf of the zombist, I concluded that we are presently unable to know whether zombies or shombies are conceivable on ideal rational reflection.

In the third and final chapter, I examined the force of Block’s (2007) overflow argument against the view that a functional analysis of experience must or will involve cognitive accessibility. I pointed out that cognitive accessibility is at least part of the conception of experience implicit in the scientific method(s) for studying experience. For this reason, without a drastic change in the scientific method(s) for studying experience, there can be no scientific case for experience that is not cognitively accessible. I noted that Block (in conversation) seems to admit that certain experiments in psychology which he takes to be evidence of experience without cognitive accessibility are not based on report or performance, but rather, on first-person phenomenological evidence. I subsequently argued that Block’s (2007) mesh argument, which involves a mesh between first-person evidence in certain psychology experiments and neuroscience, has both intuitive and empirical drawbacks. Finally, I noted that a conceivability argument against cognitive accessibility as partially constitutive of experience is ineffective because it meets with a reverse-zombie argument style counter conceivability claim.

I embrace physicalism and the functional analyzability of experience. Moreover, I think that cognitive accessibility will likely form a core part of the correct analysis of experience. Part of the reason I am sanguine about cognitive accessibility as partially constitutive of experience is that this conception of experience is implicit in our current scientific method(s) for studying experience. Another is that it is in some sense platitudinous that experience involves a kind of ‘awareness’ or ‘access:’ the notion of an experiential mental state that a subject has no awareness
of or access to is a strange one indeed. Understanding experience in terms of cognitive accessibility arguably provides a good way of capturing this platitude. Naturally, however, I’m more concerned about whether or not experience is functionally analyzable at all, than about what the most promising way of analyzing it may be. There is a deep intuitive pull in favor of the view that experience cannot be exhausted, conceptually speaking, in terms of its causal cum relational profile. It’s not obvious, however, how much evidential weight this should be accorded. I don’t have anything particularly illuminating to say about this issue, but I do hope to have shown that physicalism is a substantive metaphysical position which is not as threatened by this intuition as many seem to have thought.
Reference List


