Physicalism and the Intrinsic Nature of Consciousness

PATRICK LEWTAS *American University of Beirut*

ABSTRACT: This paper presents a metaphysical argument against physicalism based on the distinction between intrinsic and extrinsic properties. It argues that the physical, as physicalism must understand it, consists entirely of extrinsic properties, whereas consciousness involves at least some intrinsic properties. It concludes that consciousness has non-physical properties and that physicalism is false. The paper then defends its premises against current physicalist thinking. As much as possible, it offers metaphysical arguments about physical and conscious properties rather than epistemological arguments about our physical and phenomenal concepts.

RÉSUMÉ : Cet article présente, contre le physicalisme, un argument métaphysique fondé sur la distinction entre les propriétés intrinsèques et extrinsèques. Il soutient que le physique, comme le physicalisme doit le comprendre, comprend uniquement des propriétés extrinsèques, tandis que la conscience implique au moins certaines propriétés intrinsèques. Il conclut que la conscience a des propriétés non-physiques et que le physicalisme est faux. L'article défend ensuite ses prémisses contre la pensée physicaliste actuelle. Autant que possible, il offre des arguments métaphysiques portant sur les propriétés physiques et conscientes plutôt que des arguments épistémologiques portant sur nos concepts physiques et phénoménaux.

Most arguments against physicalism draw metaphysical conclusions from epistemological premises. This leaves them hostage to doubts about the epistemology. Philosophers debate the epistemology, and the metaphysics—the point, after all—too often drops from sight. This paper puts forward an

***Dialogue* 53 (2014), 203–228.**

doi:10.1017/S0012217314000237

argument from metaphysical premises to metaphysical conclusion. And it supports its premises, as much as possible, with metaphysical considerations. It can't avoid epistemological issues, of course. For that would mean talking past the physicalist. But, most places, the paper recasts these issues in metaphysical terms. It asks, for example, not how we know consciousness, but how the world must work if we know consciousness the way the physicalist says we do. And then it asks how plausible that implicit metaphysics is.

1. The Argument

The paper's argument runs as follows:

(1) The physical, as physicalism must understand it, consists entirely of extrinsic properties (relations without intrinsic relata).
(2) Conscious experience has at least some intrinsic properties.
(3) These intrinsic properties don't reduce to extrinsic properties.

Therefore:

(4) Conscious experience has properties that can't be physical, as physicalism must understand it.

Therefore:

(5) Physicalism is false.

Others make arguments, similar on the surface, which remain epistemological nonetheless. David Chalmers infers that physical accounts won't explain consciousness from premises stating that (1) physical accounts explain at most structure and function; and (2) explaining structure and function won't explain consciousness.[1] Because his premises deal with explanation rather than properties, he reaches a conclusion about the bounds of physical explanation rather than the bounds of physical nature. Gregg Rosenberg notes that science doesn't reveal categorical properties underlying physical dispositions.[2] He reasons that the physical world therefore consists either of dispositions all the way down or of dispositions resting on empirically inaccessible grounds. Physicalism can recognize these grounds, he allows, but only as empty placeholders without intrinsic content. Rosenberg then argues that neither dispositions without non-dispositional relata nor relations anchored in empty placeholders can yield the intrinsic content of conscious experience. Unfortunately he casts this last argument in terms of non-entailment. William Seager comes closest.[3] He marshals

[1] Chalmers (2003).
[2] Rosenberg (2004).
[3] Seager (2006).

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

all the elements of the argument. But he deploys them, not to attack physicalism, but to explain the explanatory gap.[4]

2. The First Premise: the Nature of the Physical

2.1 Physicalism

We need a grasp of the physical as physicalism must understand it. We therefore need to know, not only what physicalism claims, but what it must claim. Physicalism holds, very roughly, that all concrete entities are either physical or nothing over and above the physical. This takes us surprisingly far when we ask how and why physicalism makes this claim.

Physicalism does metaphysics by proxy. It delegates exploring the concrete world to the science departments down the hall. It passes the torch on concrete metaphysics because it passes the torch on epistemology. After comparing the successes of science to the failures of other approaches (including, notably, *a priori* philosophy), physicalism judges the scientific method the best way—maybe the only way—of learning about the concrete world. Overall, physicalism accepts that science has been at it long enough, and gotten good enough, that we can—and should—trust its findings, in outline at least. It follows that physicalism loses its moorings if it distances itself from the methods and findings of science. Unless its concrete metaphysic adopts the scientific image more or less as science gives it to us, physicalism becomes underspecified and indeterminate, telling us little about the world. It has specific metaphysical content, after all, only when it takes science to have gotten it largely right. And insofar as physicalism makes metaphysical claims over and above these borrowings, or relies on non-scientific methods in its own independent investigations of concrete reality, it undercuts its warrant. For that warrant rests significantly on the success of science.

Physicalism thus necessarily inherits much from science. One bequest is metaphysical reductivism. Physicalism takes on board scientists' working hypothesis, or empirical inference, that all non-basic entities are nothing over and above basic physical entities.[5] This gives it a bottom up rather than top down metaphysic. Physicalism thus accepts that all non-basic entities are (at least token) identical to composites or aggregates of basic physical entities

[4] The argument has other forerunners. Blackburn (1990) notes the dispositional nature of physical reality and the apparent categorical nature of consciousness. But he hasn't combined these points into an argument against physicalism. And Braddon-Mitchell and Jackson (2007), 131 allow that consciousness seems to have intrinsic features not captured by causal/functional accounts. But they use this to motivate the absent and inverted qualia arguments against functionalism.

[5] 'Basic' means not built of anything. Hence basic entities lack parts and stand over and above all other entities.

tokened/instanced at the bottom level.[6] It follows, for the physicalist, that all non-basic entities derive their natures from the natures of the basic physical entities that make them up. A second bequest is non-basic mentality. Physicalism draws the empirical inference that mentality is a non-basic feature of the world. It bases this on physics' silence about mentality (imprecise and disputed talk about a role for consciousness in the collapse of the quantum wave function aside). Science encounters unambiguous signs of mentality only when special sciences like biology and psychology investigate organisms with complex nervous systems. The physicalist therefore insists that consciousness results only when enough basic physical entities come together in the right ways.[7] This non-basic mentality claim has become the hallmark of physicalism as a mind-body theory. Its credo denies (1) that any mental properties have basic status; and (2) that mentality of any kind exists at the bottom level. These claims set physicalism off from views like emergentism (which understands conscious properties as basic but high-level properties) and panpsychism (which understands conscious properties, at their simplest, as basic and bottom-level properties).

Physicalism must also claim that high-level entities are (at least token) identical to arrangements of basic physical entities *as physics describes them*. Otherwise, as remarked above, physicalism undercuts its warrant and paints too indeterminate a picture of the world. It follows that physicalism must understand the physical in terms of physical theory, since nowhere else do we find the physical as physics describes it.[8] We avoid making physicalism trivially false (a risk, since current theory surely contains gaps and errors) by allowing physical theory to slide into ideal or completed theory. And we avoid emptying

[6] Most physicalists, but not all, accept a bottom level. This paper does too. Its arguments would still go through without one, however, albeit with complications. Without a bottom level, for example, basic physical properties would be those with no level below which they fail to obtain.

[7] This paper uses the results relation—results from, results in—as an umbrella catching the ways one entity can be nothing over and above other entities. The results relation covers relations like identity, composition, realization, etc.

[8] Some physicalists understand the physical as everything needed in a full account of the basic natures of paradigmatic physical objects like atoms, rocks and stars. This object-approach fails because it floats free of actual science. It thus strips physicalism of determinate content and substantive warrant. It also allows the physical to include too much (e.g. empirically inaccessible categorical grounds, if these exist, and basic conscious properties, if these collapse the quantum wave function or characterize electrons). The object-theorist can deal with consciousness by imposing a non-mentality requirement, of course (see Stoljar 2001). But this move remains *ad hoc* but for physical theory's silence about consciousness. Only an approach based on physical theory can resort to it in a principled way.

physicalism of content and making it trivially true (a risk, since ideal or completed theory could say anything) by (1) insisting on the non-mentality requirement; and (2) stipulating that any ideal or completed physics remain recognizably scientific—that it use the scientific method to collect data empirically and to explain them causally.[9]

Doctrines that do the following therefore count, for our purposes, as physicalism:

(1) they take their ontology of basic concrete entities from basic physics;
(2) they build higher-level concrete entities out of those basic physical entities;
(3) they impose a non-mentality requirement on basic entities and on the bottom level; and
(4) they allow for extensions of basic physics that (i) deal with basic concrete entities; (ii) arise through applications of the scientific method to data collection and causal explanation; and (iii) satisfy the non-mentality requirement.

This paper thus addresses what we might call 'standard physicalism,' the physicalism of the overwhelming majority of physicalists. Some physicalists endorse instead a non-standard physicalism variously called 'Kantian physicalism,'[10] 'funny physicalism,'[11] 'Type-F monism,'[12] 'Russellian monism,'[13] and 'Russellian panpsychism.'[14] Kantian physicalists accept that empirical science only reveals dispositional properties (see ahead), insist that dispositional properties have categorical grounds, and conclude that basic physical entities have an empirically inaccessible categorical nature. Most Kantian physicalists then claim that this underlying nature consists of conscious or proto-conscious properties. For five reasons this paper doesn't consider Kantian physicalism a kind of physicalism. First, Kantian physicalism violates the non-mentality condition, the centrepiece of standard physicalist philosophy of mind. Second, Kantian physicalism recognizes entities lacking empirical warrant, which not only sets it apart from standard physicalism but also denies it the scientific warrant all physicalisms claim. Third, Kantian physicalism proposes a very

9 Hempel (1980) and Crane and Mellor (1990) set out the dilemma facing the use of physical theory in a formulation of physicalism. Many physicalists propose the solutions adopted here. This paper follows Wilson (2006) especially.

10 Jackson (1998).

11 Goff (2011).

12 Chalmers (2003).

13 Chalmers (2010).

14 Chalmers (2013). Strictly speaking, Russellian panpsychism is a sub-type of Russellian monism. Russellian monism envisions conscious or proto-conscious categorical grounds. Russellian panpsychism only recognizes conscious grounds.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

different basic concrete ontology than does standard physicalism. Whereas standard physicalism (at least token) reduces all conscious properties to basic physical entities, Kantian physicalism (at least token) reduces all physical entities to basic conscious (or proto-conscious) properties. Thus Chalmers classes Kantian physicalism as a non-materialist philosophy[15] and elsewhere regards it as a kind of idealist monism.[16] Fourth, treating Kantian physicalism and standard physicalism as of a kind garbles the contrast between physicalism and panpsychism. Although some Kantian physicalists and some panpsychists share their metaphysic (hence the labels 'Kantian physicalism' and 'Russellian panpsychism'), most physicalists rightly view panpsychism as an alternative mind-body theory. Fifth, because Kantian physicalism and standard physicalism have such different metaphysics, they have very different weaknesses. Arguments against them must therefore address very different issues. This paper targets standard physicalism only.[17]

2.2 The Physical

What does physical theory say about the world? We answer this by noting what science does, can do and must do.

Science collects data through empirical means and then systematizes, interprets and explains them. Data thus come to the scientist by way of the senses.[18] They do so directly, through the world's causal effects on eyes, ears, etc. And they do so indirectly, through the world's causal effects on scientific instruments, which in turn affect eyes, ears, etc. It follows that the scientist can gather data only about dispositional properties—tendencies to affect the senses thus-and-so under such-and-such conditions.[19] The scientist then either builds

---

[15] Chalmers (2003).

[16] Chalmers (2013).

[17] See Lewtas (forthcoming a) for a critique of Russellian panpsychism.

[18] The paper uses 'data' ambiguously, although the ambiguity does little harm. Data, in the first instance, consist of external and mind-independent objects, properties, etc. which cause mental states internal to the scientist. These mental states have the form of perceptions, cognitions, etc. that the scientist takes to represent aspects of the external world. We might say that these representations preserve some of the information contained in their external causes, although talk of information broaches fraught ontological issues and is perhaps best regarded as heuristic or metaphorical.

[19] Might manifested dispositions have non-dispositional properties where dispositions themselves don't? It doesn't matter. Even if non-dispositional manifestings do occur, science could know nothing of them. The scientist can only learn about entities disposed to affect the senses. This last bottleneck winnows out any information about non-dispositional goings-on.

The paper thus assumes the falsity of direct realism. It assumes that subjects have direct access only to their own experiences. See 3.4.3 for more about this.

her model of the world directly from her data, as does a biologist who reports the discovery of a new parrot in Borneo. Or she explains her data, by postulating properties (or entities with properties) that would cause them. In the latter case, the scientist hypothesizes new dispositions, or, at most, objects construed solely in terms of new and old dispositions. She can do no other. Hypotheses framed in terms of non-dispositional properties would lack causal explanatory power and hold aloof from empirical test. They might have metaphysical warrant, but they would lack empirical warrant. The scientist, as scientist, could not posit them. The physicalist, because she takes her concrete ontology from empirical science, could not recognize them.[20]

Physics, like any science, therefore portrays the entities in its domain, which include the concrete world's basic building blocks, as systems or structures of dispositions—dispositions arranged in space and time and related one to another. Thus physics understands the electron as a more or less durable cluster or bundle of basic dispositions to do thus-and-so under such-and-such circumstances. Physics, and the rest of the sciences, with their reductive outlook, then treat non-basic entities as composites or aggregates of basic physical entities. This makes non-basic entities complex systems/structures of interacting dispositions. It follows that all concrete entities, at least as science reveals them, consist solely of structures of interacting spatio-temporally located dispositions.

Thus science and physics say nothing about any qualities which fit into or fill out physical structure. They say nothing about any qualities which underlie, hold together or ground physical dispositions. They say nothing about the intrinsic nature, if any, of physical substance—what physical substance is like in itself, apart from its powers to interact with or relate to other things.[21] As empirical disciplines they can't. Thus D.M. Armstrong's plaintive words:

> But if we look at the properties of physical objects that physicists are prepared to allow them, such as mass, electric charge, or momentum, these show a distressing tendency to dissolve into *relations* that one object has to another. What, then, are the things that have these relations to each other? Must they not have a non-relational nature if they are to sustain relations? But what is this nature? Physics does not tell us.[22]

[20] The physicalist need take only her concrete ontology from empirical science. She can, using *a priori* means more or less independent of empirical science, probe the non-concrete (or less obviously concrete) corners of her ontology: the categorial natures of properties, substances, events, causal relations, laws of nature, etc.

[21] Physicalists who come to this conclusion include Armstrong (1961, 1993), Blackburn (1990), Hume (2000) (if he counts as a physicalist), Braddon-Mitchell and Jackson (2007) (who entertain but don't definitively endorse it), Lewis (2009), Mackie (1973), Maxwell (1978), Pereboom (2011), Smart (1963) and Stoljar (2001, 2006) (although Maxwell, Pereboom and Stoljar count, for these purposes, as Kantian physicalists).

[22] Armstrong (1993), 282.

Let us use 'intrinsic properties' to refer to any properties a concrete entity has besides (1) its spatio-temporal (structural) properties; (2) its dispositional (causal/dynamical/functional) properties; and (3) any relations it bears to other entities not parts of itself.[23] Intrinsic properties cover the sorts of things about which science holds silent—qualities that fit into or fill out structures; qualities that underlie, hold together or ground dispositions; most of all, qualities an entity has in itself apart from its powers to interact with or relate to other things. Let us use 'extrinsic properties' to refer to all properties of a concrete entity other than its intrinsic properties.

3. The Second Premise: Intrinsic Conscious Nature

3.1 Experience

Conscious experience seems to have something intrinsic about it. It presents as having a determinate qualitative nature apart from its relations to, or its powers to interact with, other things. It thus appears to have a nature over and above any causal/functional properties it might also have. Braddon-Mitchell and Jackson put it this way:

> Intuitively, phenomenal nature is intrinsic. The perceived redness of a sunset is a feature of how the experience is here and now; you cannot capture its nature by talking of its similarities and differences, and of what causes it, and of what it causes. It is as intrinsic as squareness. But then perceived redness in particular, and colour experience in general, cannot be captured in functional terms, in terms of causal relations.[24]

Any experience has three aspects conceptually separable although perhaps metaphysically one: a conscious presentation (what is presented); a relation of conscious or subjective awareness (the presenting, which marks off experience from, say, the non-conscious 'perceptions' of an electric eye); and a conscious subject (what is consciously aware and thus presented to). (By 'subject' we needn't mean anything grand—just what experiences, maybe just the experiencing of the experience (think of the experienced ideas making up Hume's mind).) For two reasons we might treat these three aspects as one. First, each aspect necessitates the others. Not only can't we have an experience with one, much less two, missing. We can't even conceive of an experience without all three together.[25] Second, when we introspect carefully, we don't discern two

[23] This rules out, among other things, contrived (and 'unnatural') relational properties like being thought of by a short lady two kilometres north of a burning barn.

[24] Braddon-Mitchell & Jackson (2007), 131.

[25] Take away everything presented and you don't have experience. Take away everything presented to and you don't have experience. Take away the presenting and you don't have experience.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

distinct entities, subject and experience, standing in some quasi-perceptual relation. We notice nothing more than a stream of experiences, each of which 'gives' or 'has' all three aspects merely through being an experience.[26] If these aspects really are the same, then the relationship between experience and subject is not that of percept to perceiver, or content to container, but of part to whole.[27] And this makes sense phenomenologically. Our conscious selves, *as we experience them*, consist of all and only the experiences we have.

Experience extends widely. Philosophers may attend to qualia—experienced sensory or emotional qualities ('raw feels'). But consciousness embraces everything one has conscious awareness of. "It includes everything a bat or a new born baby can feel, and everything a great mathematician can experience in thinking."[28]

3.2 Physicalism and Experience

Physicalists divide into *a priori* physicalists—those who claim actual or possible *a priori* links between our physical and phenomenal concepts—and *a posteriori* physicalists—those who deny such links.[29]

3.2.1 A Priori *Physicalism*

We saw that physicalism necessarily understands physical nature in purely causal/functional terms. It follows that physicalism necessarily construes physical concepts as causal/functional concepts. And physicalists do construe them thus.[30]

Now, if physical concepts are causal/functional concepts, and if, as *a priori* physicalists claim, *a priori* conceptual links hold between physical and phenomenal concepts, then our phenomenal concepts must also be causal/functional concepts. This follows from the fact that *a priori* conceptual links can't hold between purely and essentially causal/functional concepts and purely and essentially non-causal/functional concepts. But if our physical concepts and phenomenal concepts are both causal/functional concepts, then the explanatory gap between consciousness and physical nature should be a conceptual impossibility (and, assuming minimal rationality on our part, a psychological impossibility as well). This explanatory gap, recall, rests on the fact that we

---

[26] Hume famously remarked upon this feature of the phenomenology.

[27] In which case we have a ready explanation for the so-called immediacy of experience. See 3.4.2 and 3.4.3.

[28] Strawson (1994), 3-4.

[29] This division lumps together as *a priori* physicalists Chalmers' (2003) Type-A and Type-C materialists. Type-C materialists concede that our current phenomenal and physical concepts lack *a priori* links. But they hold out for such links tomorrow, after some advance gives us better concepts.

[30] Kantian physicalists don't (see Maxwell, 1978). But this paper only addresses standard physicalists.

can't see, and can't see that we ever could see, how basic physical entities could result in consciousness. Now, if we already have the right determinate physical and phenomenal concepts, then, the truth of *a priori* physicalism granted, we should see today how physical entities result in consciousness. But we don't. And even if we don't have the right determinate physical and phenomenal concepts today, then, the truth of *a priori* physicalism granted, we should at least see the possibility of seeing how physical entities might result in consciousness. After all, no principled conceptual barrier can lie between the determinate of a currently unspecified causal/functional concept and the determinate of another unspecified causal/functional concept. So, if we have little or nothing to go on today but the fact that our physical and phenomenal concepts both qualify as causal/functional concepts, then it should seem at least possible that the entities referred to by the one might result in the entities referred to by the other. But it doesn't. In sum, *a priori* physicalism entails the conceptual and (assuming minimal rationality on our part) psychological impossibility of the explanatory gap. But the explanatory gap obtains nonetheless.

The *a priori* physicalist might retort that conscious properties must be causal/functional entities—and hence that our concepts of them, if adequate, must be causal/functional concepts—because consciousness must have evolved and only causal/functional entities can evolve. But the non-physicalist can concede that conscious properties evolved without granting them causal/functional natures. He can claim that they evolved as the realizers of sundry functions, which, like any realizers, have properties (including, perhaps, non-causal/functional properties) under-determined by the roles they realize.[31] More, he can hold that conscious properties, although realizers of functional roles, nevertheless have purely non-causal/functional natures. Here he might claim that conscious properties 'sit there,' altogether passively, and physical nature, of itself, actively responds to them. This puts all active causality on the physical side.[32]

---

31 See Lewis (2009).

32 This suggestion assumes that at least some causes actively produce their effects (through the exercise of powers intrinsic to their natures). It therefore rejects Humean and quasi-Humean understandings of causation. According to the Humean, intrinsic properties lack all causal efficacy but, collectively, instantiate a contingent pattern (which we can understand in terms of dependency relations like regular succession, counterfactual dependency, conditions necessary and sufficient, etc.) which our minds misinterpret as causal connections. According to the quasi-Humean, causally inert intrinsic properties have causal connections thrust upon them by contingent laws of nature (typically understood as second-order universals). These connections then take the form of dependency relations between cause and effect. The suggestion in the text rejects Humean and quasi-Humean accounts because they can't distinguish passive from active causes. For, if causation takes the form of some dependency relation (whether or not underwritten by

Most *a priori* physicalists will dismiss this reasoning as too quick. They will argue that our phenomenal concepts, appearances to the contrary, either are, or should be, causal/functional concepts. The paper deals with such claims in section 3.4.

3.2.2 A Posteriori *Physicalism*

A posteriori physicalists deny *a priori* conceptual links between our physical and phenomenal concepts. They claim that we can only know *a posteriori* the metaphysical relation—identity, realization, whatever—that holds between physical entities and conscious properties.

Now, any concept with an actual referent must be transparent, translucent, mildly opaque or radically opaque.[33] A transparent concept reveals all of its referent's essential properties. Thus your concept of a circle gives you a circle's full essence: that it consists of all points equidistant from some stationary point. A translucent concept reveals some, but not all, of its referent's essential properties. Thus your concept of a human heart reveals its functional essence (that it pumps blood) but not its physical essence (that, for example, it consists of such-and-such cells arranged thus-and-so). A mildly opaque concept reveals none of its referent's essential properties, but at least some of its accidental properties (typically enough to fix actual world reference). Thus your concept of water identifies it as the stuff filling the rivers, lakes and oceans, etc., but doesn't reveal its essential physical nature as H_2O. And a radically opaque concept reveals none of its referent's properties. Thus your concept <that stuff>, taken independently of whatever it refers to, functions as a blind demonstrative without ascriptive conceptual content. These four possibilities exhaust logical space.[34]

---

external laws of nature) that all causes bear to their effects, then so-called passive causes will count as causes in exactly the same way as so-called active causes, undercutting even the possibility of distinguishing them. But if, instead, causation issues from the natures of the causal relata, then we can distinguish different types of causation based on different kinds of causal relata. We can thus say that a cause causes actively when the causal relation (between the cause and the entity suffering the effect) issues from the nature of the cause, and that a cause causes passively when the causal relation issues from the nature of the entity suffering the effect. In both cases we allow that the causal relation, however established, then triggers some disposition in the entity suffering the effect. In passive causation, then, the entity suffering the effect, on its own, establishes the causal link with the 'cause'. And if it succeeds in doing this, it then manifests a disposition to produce the effect. At no point does the 'cause' actively do anything. See Lewtas (forthcoming b) for a theory of mental causation along these lines.

33 This taxonomy, as well as some of the arguments this section bases upon it, come from Goff (2011).

34 Allow that a concept revealing some but not all of its referent's essential properties and some of its accidental properties counts as a translucent property.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

Physicalists necessarily claim that our physical concepts are transparent or translucent. Recall that physicalism draws most of its warrant from the success of science, from the fact that science paints a picture of the world at least more or less right. The physicalist must therefore insist that our physical concepts reveal at least some (and likely most if not all) of their referents' essential properties.

What about our phenomenal concepts? The *a posteriori* physicalist can't allow that these are transparent. For suppose experience E has essential properties A, B, C and D. Then our phenomenal concept of E, $C_{PHEN}(E)$, would represent E as having A, B, C and D. More, since it truly represents the essence of E, it represents E as having A as A truly is (A *qua* A, rather than, say, A *qua* Josie's favourite quality), as having B as B truly is, and so on. The *a posteriori* physicalist therefore can't say that we have two conceptually independent grasps of each of A, B, C and D. $C_{PHEN}(E)$ represents A, B, C and D as, respectively, A, B, C and D. Otherwise it fails to reveal E's essential properties, contrary to hypothesis. But the *a posteriori* physicalist also insists that each of A, B, C and D is a physical property. It follows that $C_{PHEN}(E)$ represents E as having the very physical properties it does have. But then it should be *a priori* that E is physical, contrary to *a posteriori* physicalism.

The *a posteriori* physicalist can't allow that our phenomenal concepts are translucent. For suppose, again, that experience E has essential properties A, B, C and D. Suppose, also, that $C_{PHEN}(E)$, being translucent, represents E as having A and B. But, as above, $C_{PHEN}(E)$ represents E as having A as A truly is and as having B as B truly is. And A and B are physical properties. It follows that $C_{PHEN}(E)$ represents E as an entity having only physical properties (albeit not all the physical properties it in fact has). But then, as above, it should be *a priori* that E is physical, contrary to *a posteriori* physicalism.

The *a posteriori* physicalist can't allow that our phenomenal concepts are mildly opaque. Suppose that experience E has essential properties A, B and C and accidental properties d and e. And suppose that our phenomenal concept of E, $C_{PHEN}(E)$, represents E as having d and e. Since d and e are only accidental properties of E, E could lack them and still be the kind of experience it is. But it isn't plausible that our phenomenal concepts say nothing about the kinds of things they refer to. For instance, it seems that the experience of pain can't lack all the properties our concept attributes to it, including felt hurtfulness, and still count as pain.[35] It seems that our phenomenal concepts must reveal more than mildly opaque concepts can.

The *a posteriori* physicalist can't allow that our phenomenal concepts are radically opaque. In this case our concept of experience E, $C_{PHEN}(E)$, doesn't represent E as having any particular properties. But then $C_{PHEN}(E)$ tells us nothing whatsoever about E—in particular, nothing whatsoever about what it

[35] Kripke (1980), 144-155 argues thus.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

is like to have E. But, again, it seems implausible that our concept of the experience of pain, which attributes felt hurtfulness, tells us nothing whatsoever about the experience it refers to.

It seems, in sum, that our phenomenal concepts must be either transparent, translucent, mildly opaque or radically opaque, since these four exhaust the options. But none works for *a posteriori* physicalism. This bodes ill for *a posteriori* physicalism.

Note that the above arguments attend to properties had and properties referred to. They thereby focus, as much as their target allows, on the metaphysics presupposed by *a posteriori* physicalist epistemology. The arguments proceed, not by directly challenging *a posteriori* physicalist epistemology, but rather by drawing untoward consequences out of its implicit metaphysics. The *a posteriori* physicalist, for his part, can resist these arguments in two ways. He can make the epistemological claim that, appearances and intuitions to the contrary, our phenomenal concepts so badly misrepresent their referents as to reveal nothing, or next to nothing, about them. In other words, he can maintain that our phenomenal concepts are mildly or radically opaque notwithstanding the arguments above. Section 3.4 addresses arguments to this effect. Or he can make the metaphysical claim that the world possesses modal structure that denies our phenomenal concepts any solid grip upon it. The next section looks at this.

3.3 Brute Metaphysical Identities

Some *a posteriori* physicalists postulate brute identities or brute (or strong[36]) metaphysical necessities. These come down to the same thing, seen as identities when viewed metaphysically and seen as necessities when viewed modally. If A and B stand in a relationship of brute identity, then A equals B, but nothing—nothing metaphysically real—grounds an epistemically intelligible connection between them. Thus the non-identity of A and B won't seem impossible no matter how much we know about A and B. By the same token, if A and B instance a brute metaphysical necessity, then their non-identity remains logically possible, and therefore *a priori* conceivable, but is nonetheless not metaphysically possible.

Brute identities/necessities have many shortcomings.[37] Here we focus on their metaphysical incoherence. We therefore view them as identities, which

[36] Different philosophers use different terms. Chalmers (1996, 2003, 2010) speaks of strong metaphysical necessities. Levine (2001) speaks of brute metaphysical necessities (partly because he distinguishes strong metaphysical necessities, which he deems (barely) acceptable, from brute metaphysical necessities, which he deems unacceptable).

[37] Chalmers (2010), 166-192 offers many arguments against brute metaphysical necessities based on (what he sees as) their modal profligacy and the irrationality of postulating them. He presents on p.170 a brief list of these concerns.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

allows us to think of properties and the relations between properties, rather than as necessities, which, as modal entities, steer our thinking toward conceivability, rationality, and other similarly epistemological or quasi-epistemological phenomena of uncertain ontological status. So we ask how property A (a conscious property) and property B (a physical property) can stand in a relationship of identity where nothing about A and nothing about B make this intelligible—and yet where, by hypothesis, this unintelligibility results from the metaphysics and not merely from epistemology. If A and B really are the same thing—if we have just one thing rather than two distinct things somehow related—then we have a single thing with a single set of properties. Now, if the apparent bruteness doesn't result from epistemological factors, then we don't cognize the single property A/B as property A through one subset of its properties and as property B through another and (wholly or partly) disjoint subset of its properties. Nor do we cognize A/B as A through a subset of its properties plus some set S_1 of properties it doesn't in fact have, and as B through another (possibly identical) subset of its properties plus a different set S_2 of properties it doesn't in fact have. We instead cognize A/B as A and as B through cognizing the very same properties. More, we cognize A and B in the very same way, since the bruteness, again by hypothesis, doesn't spring from epistemological factors. But then how can the identity or necessity be brute—and therefore unintelligible in principle? How can it fail to be straightforwardly intelligible?

In sum, brute necessities are metaphysical posits that perform two roles in *a posteriori* physicalist thinking. They ensure the tightest possible metaphysical link between conscious and physical properties: identity. And they account for our inability to grasp this link. But when we ask about the metaphysics presupposed by the epistemological role the *a posteriori* physicalist would have brute identities play, we discover that no single posit could fulfil both of the demands the physicalist places upon it.

3.4 Knowledge of Experience

Given the above arguments, the physicalist, *a priori* or *a posteriori*, must say that experience is not what it seems. Somehow our concepts and/or representations get it wrong.

3.4.1 Experience's Hidden Nature?

Some philosophers claim that experience has hidden depths inaccessible to experience. We would know experience as physical if only we could access these depths. Thus Armstrong:

> I would…advance the idea…that everything in the world, *everything*, every event, every property of things and events, every relation that things and events have to each other, are each one of them an epistemological *iceberg*. Our knowledge and rational belief about all these things, though real, is selective and limited. If you take

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

> this view then it becomes much easier to accept that the secondary qualities might have hidden depths to which we cannot penetrate in perception.[38]

This makes sense only if experience includes more than we experience. But it doesn't, of course. We can't define a thing to include something other than itself. We certainly can't define a thing to include something it isn't. The thing we wish to understand is experience itself—the property of experience *as experienced*. So we define experience that way. If that property has hidden depths, if its roots reach into the unexperienced, then we acknowledge that experience relates to non-experiential entities, perhaps even by way of a results-from/results-in relation. But if, at the outset, we define experience to include the unexperienced, then we change the subject, from experience as experienced to an amalgam of experience and something else. We thereby risk explaining something other than what we care about: we risk explaining the unexperienced something else instead of experience itself.

Derk Pereboom offers a variant of this approach.[39] He argues that experience is straightforwardly physical. However, its mode of presentation, the way it presents to us, makes experience appear to have properties that in fact it lacks. And here Pereboom has in mind the very 'phenomenal properties' that, when we compare them to physical properties, give rise to the explanatory gap. But this move fares no better than Armstrong's. If Pereboom has it right, we don't experience experience, but instead this other thing, its mode of presentation. But then Pereboom's 'mode of presentation' is nothing other than experience as we understand it, and his 'experience' is something else, something other than experience, something we don't in fact experience. Pereboom must now say either that this mode of presentation has the phenomenal properties in question, in which case his re-labelling hasn't advanced his cause. Or he must deny that any entity has them. But in that case he eliminates experience altogether, along with the phenomenal properties that make it so difficult to understand in physical terms. But this leaves it baffling why the explanatory gap was a problem in the first place—why Pereboom went to such trouble to explain it away.

In sum, philosophers like Armstrong and Pereboom argue for a relation between experience and something else. Then they explain the something else. In Armstrong's case, the physicalist understands the something else as experience's hidden and unexperienced essence. In Pereboom's case, the physicalist re-labels the experience as the something else and the something else as the experience—switching their labels, as it were—and then attends

[38] Armstrong (1999), 129-130. Armstrong counts as a higher-order perception theorist: he believes consciousness results from a perception-like awareness of other mental states. So "perception," at the end of the passage, means experience.

[39] Pereboom (2011).

to and physically explains the now unexperienced 'experience'. In both cases the physicalist gets the epistemology wrong because he makes a basic metaphysical mistake.

3.4.2 The Immediacy of Experience

Most physicalists deny unmediated access to experience. They claim we have no better access to experience than to external objects represented in experience. They must say this. Otherwise they couldn't easily maintain that our phenomenal concepts/representations get it wrong. And then they would find themselves pressed to take the explanatory gap at its face.

The denial of immediacy fails, however. The physicalist must say that experience involves a mediated experience of an experience—where E_1, the experience the subject is conscious of, gives mediated access to E_2, some underlying but nonetheless experiential state. For, if the state E_2, access to which is mediated, isn't an experience, then the physicalist doesn't rebut immediate access to experience. And if the state doing the mediation, E_1, isn't an experience, then the physicalist denies the subject experience, eliminating experience instead of reducing it to the physical. But how do E_1 and E_2 relate? Note that E_1, an experience, can give mediated access to E_2, another experience, in such a way that the subject's overall experience is a mediated experience, only if (1) the subject experiences *as an experience* E_1's representation of E_2; and (2) this experiential representation of E_2 is what the subject experiences. This has the result that E_1 is a reproduction, facsimile or likeness of E_2 (E_1 can't do what the physicalist asks and merely refer to E_2 or otherwise represent E_2 symbolically). It follows that E_1 faithfully represents E_2 only if it shares E_2's properties. But now we see that E_1 can't be the same state as E_2. For E_1, by hypothesis, misrepresents E_2 (the physicalist, after all, denies immediacy only to claim that experience is not what it appears), and no state can be identical to an inexact reproduction of itself. (If E_1 is an inexact reproduction of E_2, then it represents E_2 as other than E_2 really is—as having properties E_2 doesn't have or as lacking properties E_2 does have. But then E_1, the inexact reproduction, can't share all its properties with E_2, the state it is an inexact reproduction of. It must be distinct from E_2.) But if E_1 is distinct from E_2, and E_1 is what the subject experiences, then it doesn't help the physicalist that E_1 gives mediated access to E_2. If all experience is mediated, then the denier of immediacy needs somehow to mediate the subject's access to E_1. But he can do this only by introducing an E_3, between the subject and E_1, that does mediate E_1. And this E_3 must itself count as an experience (for otherwise the subject experiences nothing, again contrary to hypothesis). But now the physicalist needs another experience, E_4, to mediate E_3, and so on. The vicious regress ends only if the physicalist concedes some experience had immediately.

Here we've looked at the metaphysics implicit in the physicalist's epistemological denial of immediacy. And we've found that this metaphysics involves either the existence of immediacy or the non-existence of experience. But neither of these coheres with non-eliminative physicalism.[40]

3.4.3 Self-presentation and Self-revelation

It follows from immediacy that experience can't support a distinction between reality and appearance. The experience, as it really is, and as it appears to the subject, are necessarily one and the same. This makes good sense if, as suggested in 3.1, the subject relates to his experience as whole to part rather than as perceiver to percept. In this light, we speak loosely when we say the subject has direct access to his experiences. We speak loosely even when we say he encounters them through basic acquaintance. In both cases we risk the obscure claim that the subject exists separately from his experiences while nevertheless having immediate epistemic access to them. Treating the subject as the sum of his experiences dispels the mystery. For now we see the immediacy of experience as a straightforward consequence of the metaphysics of identity.

The non-physicalist can nonetheless both acknowledge and explain our misjudgements about our conscious states. Even though experience is unmediated, our judgements, beliefs and knowledge about it are mediated. These issue from cognitive processes that take experience as input and generate, as output, judgements, beliefs and knowledge about experience. And these cognitive processes, like all cognitive processes, can err. This opens up the possibility, at the higher-order cognitive level, of a gap between appearance and reality—of a discrepancy between the first-order experience and the way the higher-order state represents it. We can sum this up by saying that experience always, and necessarily, self-presents (at the level of first-order experience), but doesn't necessarily self-reveal (at the higher-order level involving representations of experience).[41]

---

[40] See section 4 for very brief remarks about elimativism about consciousness.

[41] Christopher Hill (1991), 128 comes to the same conclusion. He contrasts errors of ignorance with errors of judgement. Errors of ignorance occur when we base our beliefs on imperfect information. Errors of judgement occur when we fail to take into account all the information we have. Hill argues that we can't make errors of ignorance when forming beliefs about experience—because we have direct access to the full experience. But nothing stops us from making errors of judgement. Heil (2003), 237-238 draws a similar conclusion.

The claims in the text don't rule out conceptual experiences or conceptual seemings. For instance, where a subject experiences something as thus-and-so (e.g., when I experience a dog *as* a dog), the conceptual element lies inside the experience as one of its elements. And the subject can have this experience as immediately as any other. Conceptual experiences and seemings become troublesome only where an experience E_1 takes another experience E_2 as thus-and-so. Here the subject immediately experiences E_1, and thus also E_1's representation of E_2. But this representation

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

This paper's arguments rely, of course, on judgements and knowledge claims about experience, and not just on experience itself. But the fact that experience self-presents places tight constraints on how badly it can fail to self-reveal. If the full experience, as it is, forms part of the conscious mind, then that mind is well-placed to know and judge it. Not surprisingly, then, most of our misjudgements about experiences happen at the margins. They typically concern very short experiences, subtle emotional discriminations, unflattering self-attributions, etc.[42] But the paper's argument doesn't turn on anything rare or fine. It requires only that experience have at least one intrinsic property. Yet every experience presents as having intrinsic properties, even very short experiences, subtle emotional states difficult to label, and unflattering self-attributions. Even conscious doubts about the intrinsic nature of experience themselves present with it. It follows that the physicalist, to undercut the argument, must cast doubt on the representational fidelity of experience *per se*.

But he can't easily do this. Because the physicalist borrows most of his concrete metaphysic from science, he must accept empirical science as a valid way of knowing the world. He must also accept the picture painted by empirical science as a more or less complete and accurate representation of the world. But almost all scientific data reach (the cognitive domains of) our minds through consciousness. More, our explanatory theories tell us about the world only when they figure in consciousness. If we misjudge our experiences as badly as the physicalist needs, then, it would seem, we can trust none of the representations presented therein. But then we must distrust all empirical data and all theories built upon them. This means rejecting both the scientific method as a valid way of exploring nature and the scientific understanding of the world as our best model of it. But then the warrant for physicalism crumbles. The physicalist must therefore concede that self-presentation entails enough self-revelation to ground empirical science. But we need much less self-revelation than this to judge that experience involves intrinsic qualities. Whereas physicalism needs conscious representations reliable enough to accurately convey scientific information about external entities, this paper's argument

doesn't count as an immediate experience of E_2 (we must assume, of course, that the subject once experienced E_2 immediately; otherwise the subject wouldn't have E_2 to represent). For E_1's representation of E_2 involves an interpretation of E_2 generated by some cognitive mechanism. The important point nevertheless stands: the paper's talk of immediacy and its distinction between self-presentation and self-revelation doesn't rule out conceptual experiences or conceptual seemings. We can immediately experience thoughts, beliefs, opinions, reasonings, understandings, confusions, etc. Recall Strawson's words: "[Experience] includes...everything a great mathematician can experience in thinking" (Strawson, 1994, 3-4).

[42] See Churchland (1999), 78 and Heil (2003), 237-238.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

only needs minimal knowledge of experience as presented. And scientific information, theoretical and empirical, is so thoroughly enmeshed in experience that the physicalist can't plausibly claim that experience perfectly reveals its informational content while wholly obscuring the experiential form that content takes.[43]

You might object that this shows, at most, that the physicalist can't say that the first person perspective massively misrepresents, but not that it doesn't in fact do so. Bear in mind, however, where the burden of proof lies. Anyone who would know and understand the world must start with appearances. More, she must take appearances at face value barring specific, positive and persuasive reasons to the contrary. If the would-be theorist turns her back on appearances as such, then her project of learning and understanding could never get started. And, once started, it could never settle on a fixed picture or model of the world. Now, we have every reason to take the first person perspective at face value but for considerations put forward by the physicalist. The burden therefore lies on her to discredit its representational fidelity.

But we have good *metaphysical* reasons for thinking this burden very hard to meet. Earlier, the paper voiced its rejection of direct realism.[44] Before that it ventured that conscious subjects consist of their experiences. A few pages ago it argued for the immediacy of experience. These points together call to mind the following picture (sketched here in roughest outline and intended more to suggest than to 'prove'). Conscious subjects evolved to use their experiences to represent the external world. They therefore have direct access only to an internally generated model of the world. This model serves their needs because, in adaptively relevant ways, it has features isomorphic to the world (and thus supplies information about the world). It stands to reason, then, that subjects naturally and unreflectively take (at least many of) their experiences *as* the external world. This nicely explains the so-called transparency of experience that philosophers like Dretske, Harman and Tye make so much of.[45] But in this case the properties of experience don't collapse into the properties of the external world represented in experience. Instead, experience has, through identity, the very properties with which it represents the external world, and thus, in many cases at least, the very properties it represents the external world as having. From

[43] Several physicalists, when explaining why experience doesn't seem physical, have suggested that phenomenal concepts have a quotational structure incorporating the experience itself (or a fainter copy or 'replay' of it) into the concept (see, for example, Block 2007 and Papineau 2004, 2007). These physicalists can't easily maintain that our representations of experience distort or hide its nature.

The paper very shortly addresses the retort that experience, by its nature, represents the external world well while representing itself poorly.

[44] See note 19.

[45] Dretske (1997), Harman (1990) and Tye (1995).

this point of view we should expect experience to do a *better* job revealing its own nature than the nature of the external world. It therefore shouldn't surprise us that experience often represents the world as having qualities—such as the so-called secondary qualities—that we have trouble understanding the external world as actually having (except in the form of simulacra like dispositions to cause such-and-such experiences in us or the often disjunctive causal bases of those dispositions).

We can strengthen the point further. While experience serves mostly to represent the external world, it also monitors itself.[46] Presumably introspection evolved to help it do this. Thus we have experiential states, like pain, that may signify external conditions (such as tissue damage), and thereby alert the subject to them, but that don't obviously or straightforwardly represent those external conditions.[47] In this light it makes little sense to insist that such experiences—again, think of pain—better reveal the nature of (aspects of) the external world than their own experiential natures. Here claims of transparency appear weak. And it seems just as strained to insist that our conceptions of those experiences hide or massively distort their natures. Here the paper's argument that self-presentation guarantees near perfect self-revelation appears strong. After all, does our conception of pain really entitle us to say that the experience of pain—pain as we feel it—has no intrinsic features?

The physicalist can nevertheless tough it out. She can insist that our conceptions of experience do hide or massively distort its nature and that any contrary arguments beg the question. She doesn't thereby contradict herself. But this paper has shown, it hopes, how forced and *ad hoc* this move is, how little drives it but the will to uphold physicalism come what may.

3.5 A Posteriori Physicalism Misses the Mark

Most *a posteriori* physicalists purport to explain why our phenomenal concepts fail to represent experience as physical. But they neglect to explain away the fact that we both experience and represent experience as having features (at least apparently) over and above the physical. We experience and represent it, that is, not merely as not physical, but also as (at least apparently) non-physical—as having intrinsic non-causal/functional properties that physical entities, understood in causal/functional terms, can't have. So most *a posteriori* strategies, at the end of the day, fail to address the full problem of consciousness.[48]

---

[46] Higher order perception and higher order thought accounts focus on this aspect or function of consciousness.

[47] Molnar (2003), 74-80 argues, for instance, that pain doesn't represent bodily damage. However, it informs us about bodily damage because it naturally means bodily damage (in Grice's 1989 sense of natural meaning).

[48] Stephen White's property dualism argument, like the one in the text, builds on the fact that we experience experience, not only as not physical, but also as non-physical. Furthermore, and again like the argument in the text, White's argument marshals this fact against *a posteriori* physicalism. See White (1986, 2007, 2010).

We can bring this out vividly. Let feature X be the intrinsic feature (or features) of experience (at least apparently) over and above the physical—the feature that underlies the explanatory gap. The *a posteriori* physicalist admits he can't explain how merely physical entities come together to yield feature X. So he turns instead to our concept of X. But this concept includes a representation as of X. The *a posteriori* physicalist's gambit therefore succeeds only if he can physically explain this representation. But how can physical entities yield a representation as of X when they can't yield X itself? The difficulty stands out starkly when we recall two things: (1) that experience doesn't support a distinction between appearance and reality; and (2) that self-presentation entails substantial, perhaps near perfect, self-revelation. It follows that the representation as of X (as provided by the phenomenal concept) and X itself (as given in consciousness) are either the same thing or very nearly so, and certainly the same kind of thing. The physicalist therefore physically explains the phenomenal concept only if he physically explains consciousness itself. But, if he could do that, he wouldn't have bothered with phenomenal concepts in the first place.[49]

Here, again, a close look at the implied metaphysics uncovers problems with the physicalist's epistemology.

4. The Third Premise: Non-reducibility of Intrinsic Nature

Section 3 establishes that experience has or involves intrinsic nature, or at least that the physicalist gives us no good reason not to take at face value the appearance that it does. The physicalist must therefore show that this intrinsic nature in fact reduces to purely relational or extrinsic nature—that each intrinsic property of experience is (at least token) identical to purely relational properties.

Daniel Stoljar begins a case to this effect.

> [O]ne can derive the instantiation of an intrinsic property from a relational one just by shifting what thing you are talking about. For example, being a husband is a relational property of Jack Spratt, and being a wife is a relational property of his wife. But being married is an intrinsic property of the pair (or the sum) of Jack Spratt and his wife. To take a different example, it seems plausible to say that I have the property of having a hand intrinsically, but my having this property obviously follows from a relation between my hand and the rest of my body, and that the truth concerning this is a relational truth.[50]

[49] Chalmers (2007) uses conceivability arguments to reach a similar conclusion: any explanation of phenomenal concepts true enough to experience to capture their nature would grant too much phenomenality to allow for a physicalistic explanation of them.

[50] Stoljar (2006), 152.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

Pereboom resists this reasoning by distinguishing absolutely purely intrinsic properties from comparatively intrinsic properties.[51] Whereas each of an entity's comparatively intrinsic properties reduces to extrinsic properties of that entity's parts, none of an entity's absolutely purely intrinsic properties does.[52] Properties like having a given shape, having a hand and being a married couple count as comparatively intrinsic properties of their instantiating entities. But which kind of property is consciousness?[53]

Maybe a quick answer lies at hand? After all, we seem to contradict ourselves when we deny that consciousness counts as, or at least involves, an absolutely purely intrinsic property. Recall that section 2 defined 'intrinsic properties' to cover properties other than spatio-temporal (structural) properties, dispositional (dynamical and/or causal/functional) properties, and all other relational properties. Section 3 then showed that conscious properties have this kind of intrinsic nature. And section 2 showed that physical entities, as physicalism must understand them, don't and can't. The paper thus framed its definitions so as to make it hard to see how the intrinsic nature of consciousness could reduce to the extrinsic nature of physical reality. (This result doesn't issue from definitional sleight of hand. It took honest work to conclude that consciousness and physical nature satisfy the respective definitions.) We see the contradiction, more clearly perhaps, if we set aside intrinsic and extrinsic properties and attend instead to the absurdity of identifying purely non-causal/functional properties (which characterize consciousness) with purely causal/functional properties (which exhaust physical nature).

The physicalist might still argue that consciousness, while admittedly apparently absolutely purely intrinsic, nevertheless in fact reduces to extrinsic (and hence physical) elements. But we saw in section 2 that basic physical dispositions, combined, merely yield higher-level dispositional structures, and that causal dynamics, combined, merely yield more of the same. We might then ask how intrinsic nature breaks down into extrinsic elements if combinations of extrinsic elements seem incapable of resulting in intrinsic nature.

The physicalist might dig in his heels. He might argue that intrinsic nature does so reduce even though we don't see how, even though we don't see how it ever could. But here he makes the same kinds of epistemological moves treated in section 3. More, these moves, because they apply to conscious and physical entities just like their section 3 analogues, presuppose the same kinds of underlying metaphysics. The same conclusion therefore holds: we have good metaphysical reasons to dismiss these moves.

The physicalist might dig in his heels even more resolutely. He might deny the existence of intrinsic or qualitative nature altogether, as do those who embrace

[51] Pereboom (2011).

[52] Op. cit., 93-94.

[53] Oddly, Pereboom (2011) says little about this.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

a purely power-based or dispositional view of reality (dispositions all the way down) and those extreme structuralists who understand nature as structures ungrounded in any intrinsic or qualitative reality (free-standing structures not instantiated by anything—dances without dancers).[54] This isn't the place to argue at length against the powers conception or extreme metaphysical structuralism. Suffice it to say, however, that each such metaphysic, by denying intrinsic, qualitative nature, either denies consciousness outright, or else diagnoses it as an illusion, as a misrepresentation of purely dispositional structure. But section 3 showed that consciousness can't sustain a distinction between reality and appearance. This rules out consciousness as an illusion of something else. The powers-view and extreme structuralism therefore imply, at the end of the day, the elimination of consciousness. Again, this is no place to tackle eliminativism about consciousness.[55] The paper instead takes the existence of experience for granted and treats its denial as a reductio of any doctrine with elimination as a consequence.

Where does this leave us? On its face, it seems implausible that intrinsic nature could reduce to purely extrinsic nature. It seems implausible that qualities, at bottom, are nothing but relations. It seems implausible that the world contains structure and nothing but. The physicalist who argues otherwise must tread one of two paths. He must make epistemological claims whose metaphysical presuppositions section 3 saw fit to reject. Or he must endorse a metaphysics implying the non-existence of consciousness itself. Yet it seems that we have far better reason to believe in consciousness than in any philosophical doctrine, much less one incompatible with consciousness. Indeed, if, as the paper suggests, we, as subjects, consist of our conscious experiences, then we can have no good reason to entertain a theory which denies the existence of consciousness.

5. Conclusion

Section 1 set out the paper's metaphysical argument against physicalism. Sections 2 through 4 gave support, respectively, to each of the argument's three premises. These sections sought, wherever possible, to adduce metaphysical rather than merely epistemological considerations. In this the paper departs from standard anti-physicalist approaches current in the literature. Now, if this support holds up, then the argument's sub-conclusion, that conscious experience has properties beyond the physical, follows straightforwardly. As does the conclusion, that physicalism is false. This still leaves open which metaphysical theory does the world justice where physicalism doesn't. The paper, through its arguments and sub-conclusions, hopes, however, to have laid down some metaphysical constraints on any such theory.

[54] See, respectively, Harré and Madden (1975) and Ladyman, Ross, Spurrett and Collier (2009).

[55] See Levine (2001) for a lengthy (and forbearing) critique of eliminativism.

References

Alter, T. and S. Walter.
2007 *Phenomenal Concepts and Phenomenal Knowledge*. New York: Oxford University Press.

Armstrong, D.
1961 *Perception and the Physical World*. London: Routledge and Keegan Paul.

Armstrong, D.
1993 *A Materialist Theory of the Mind, Rev'd Ed*. New York: Routledge.

Armstrong, D.
1999 *The Mind-Body Problem: An Opinionated Introduction*. Boulder: Westview Press.

Blackburn, S.
1990 Filling in Space. *Analysis*. 52: 60–65.

Block, N.
2007 Max Black's Objection to Mind-Body Identity. In T. Alter and S. Walter.

Braddon-Mitchell, D. and F. Jackson.
2007 *Philosophy of Mind and Cognition*. Oxford: Blackwell.

Chalmers, D.
1996 *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Chalmers, D.
2003 Consciousness and Its Place in Nature. In S. Stich and T. Warfield *Philosophy of Mind*. Oxford: Blackwell: 102–142.

Chalmers, D.
2007 Phenomenal Concepts and the Explanatory Gap. In T. Alter and S. Walter.

Chalmers, D.
2010 *The Character of Consciousness*. New York: Oxford University Press.

Chalmers, D.
2013 Panpsychism and Panprotopsychism. *The Amherst Lecture in Philosophy*. 8: 1 – 35. <http://www.amherstlecture.org/chalmers2013/>.

Churchland, P.
1999 *Matter and Consciousness*. Cambridge: MIT Press.

Crane, T. and H. Mellor.
1990 There is No Question About Physicalism. *Mind*. 99: 185–206.

Dretske, F.
1997 *Naturalizing the Mind*. Cambridge: MIT Press.

Goff, P.
2011 *A Posteriori* Physicalists Get Our Phenomenal Concepts Wrong. *Australasian Journal of Philosophy*. 89/2: 191–209.

Grice, P.
1989 *Studies in the Way of Words*. Cambridge: Harvard University Press.

https://doi.org/10.1017/S0012217314000237 Published online by Cambridge University Press

Harman, G.

1990 The Intrinsic Quality of Experience. In *Philosophical Perspectives*, 4. Oxford: Blackwell.

Harré, R. and E. Madden.

1975 *Causal Powers: Theory of Natural Necessity*. Malden: Blackwell Publishers.

Heil, J.

2003 *From an Ontological Point of View*. New York: Oxford University Press.

Hempel, C.

1980 Comments on Goodman's Ways of Worldmaking. *Synthese*. 45: 193–199.

Hill, C.

1991 *Sensations, A Defense of Type Materialism*. New York: Cambridge University Press.

Hume, D.

2000 *A Treatise on Human Nature*. D. Norton and M. Norton, eds., New York: Oxford University Press.

Jackson, F.

1998 *From Metaphysics to Ethics*. New York: Oxford University Press.

Kripke, S.

1980 *Naming and Necessity*. Cambridge: Harvard University Press.

Ladyman, J., D. Ross, D. Spurrett and J. Collier.

2009 *Every Thing Must Go: Metaphysics Naturalized*. New York: Oxford University Press.

Levine, J.

2001 Purple Haze. New York: Oxford University Press.

Lewis, D.

2009 Ramseyan Humility. In D. Braddon-Mitchell and R. Nola *Conceptual Analysis and Philosophical Naturalism*. Cambridge: MIT Press.

Lewtas, P.

Forthcoming a Russellian Panpsychism: Too Good to be True? unpublished ms.

Forthcoming b How to Make Dualist Interactionism Work. unpublished ms.

Lycan, W.

2009 Giving Dualism its Due. *Australasian Journal of Philosophy*. 87 (4): 551–563.

Mackie, J.L.

1973 *Truth, Probability and Paradox: Studies in Philosophical Logic*. Oxford: Clarenden Press.

Maxwell, G.

1978 Rigid Designators and Mind-Brain Identity. In C. Savage. *Minnesota Studies in the Philosophy of Science*, Vol. IX, Minneapolis: University of Minnesota Press.

Molnar, G.

2003 *Powers: A Study in Metaphysics*. New York: Oxford University Press.

Papineau, D.
2004 *Thinking about Consciousness*. New York: Oxford University Press.

Papineau, D.
2007 Phenomenal and Perceptual Concepts. In T. Alter and S. Walter.

Pereboom, D.
2011 *Consciousness and the Prospects of Physicalism*. New York: Oxford University Press.

Rosenberg, G.
2004 *A Place for Consciousness: Probing the Deep Structure of the Natural World*. New York: Oxford University Press.

Seager, W.
2006 The "Intrinsic Nature" Argument for Panpsychism. In G. Strawson, *et al*. *Consciousness and its Place in Nature*. Exeter: Imprint Academic.

Smart, J.J.C.
1963 *Philosophy and Scientific Realism*. London: Routledge and Keegan Paul.

Stoljar, D.
2001 Two Conceptions of the Physical. *Philosophy and Phenomenological Research*. 62: 253–281.

Stoljar, D.
2006 *Ignorance and Imagination: on the Epistemic Origin of the Problem of Consciousness*. New York: Oxford University Press.

Strawson, G.
1994 *Mental Reality*. Cambridge MA: MIT Press.

Tye, M.
1995 *Ten Problems of Consciousness*. Cambridge: MIT Press.

White, S.
1986 Curse of the Qualia. *Synthese*. 68/2: 333–368.

White, S.
2007 Property Dualism, Phenomenal Concepts, and the Semantic Premise. In T. Alter and S. Walter.

White, S.
2010 The Property Dualism Argument. In R. Koons and G. Bealer. *The Waning of Materialism*. New York: Oxford University Press.

Wilson, J.
2006 On Characterizing the Physical. *Philosophical Studies*. 131: 61–99.