

The Irrationality of Physicalism

Pat Lewtas

Received: 15 September 2013 / Accepted: 3 December 2013 / Published online: 15 December 2013
© Springer Science+Business Media Dordrecht 2013

Abstract This paper argues, not that physicalism is wrong, but that it is irrational. The paper defines standards of rationality, both metaphysical and epistemological, that physicalism necessarily inherits from science. Then it assesses physicalist efforts to naturalize consciousness in light of these. It concludes that physicalism allows its metaphysics to outrun its epistemology, in defiance of applicable standards, revealing a fundamental incoherence in the doctrine. The paper also briefly reviews other naturalization programs, to claim that physicalism, unlike the sciences, hasn't proved fruitful.

Keywords Physicalism · Rationality · Epistemic norms · Consciousness · Naturalization programs

Consider three important mind–body theories: physicalism, emergentism and panpsychism. Physicalism holds that all concrete entities are either basic physical entities or nothing over and above basic physical entities.^{1,2} This means that all higher-level concrete entities, including conscious properties, (at least token) reduce to basic physical entities at the bottom level. Every kind of physicalism cleaves, one way or another, to this core claim.³ Emergentism holds that all higher-level concrete objects and most higher-level concrete properties work this way. However, some

¹ This paper uses “entity” as a label for an existent of any kind. Objects, properties, events, processes, causes, etc. all count as entities.

² A basic entity is a basic building block not composed of anything else. Non-basic entities are composed of basic entities.

³ See Lewtas (forthcoming a) for more about the doctrine, including a more detailed formulation of it.

P. Lewtas (✉)

Department of Philosophy, American University of Beirut, Beirut, Lebanon
e-mail: pl03@aub.edu.lb

higher-level concrete properties, such as conscious properties, don't reduce to bottom-level entities. Although complex, these higher-level properties nevertheless count as basic entities. They emerge—they pop into existence out of nothing (without being built of anything, much less anything pre-existing)—upon the tokening of suitably structured non-basic physical entities.⁴ Panpsychism holds, with physicalism, that all higher-level concrete entities (at least token) reduce to basic bottom-level entities. And it denies, with emergentism, that conscious properties reduce to anything non-conscious. It therefore claims that higher-level conscious properties (at least token) reduce to basic conscious properties at the bottom level. The most reasonable kind of panpsychism attributes these basic conscious properties to basic physical objects.⁵

Focus for a moment on panpsychism. Panpsychism issues from a steadfastly rationalist outlook, rooted in (at least) a moderately strong principle of sufficient reason applied across the board. Panpsychism therefore spurns both metaphysical and epistemological gaps. It insists that all higher-level entities reduce, and reduce intelligibly, to bottom-level entities. It parts with emergentism because emergentism allows higher-level entities that don't reduce to bottom-level entities. And it parts with physicalism because, as panpsychists see it, physicalism allows higher-level entities that reduce, but not intelligibly, to bottom-level entities.⁶ Because panpsychism holds metaphysics and epistemology to the same reductive standard, it keeps metaphysical and epistemological explanations from drifting apart. In each case a good explanation shows how higher-level entities are either composites, or straightforward results, of basic bottom-level entities. (This paper uses the results relation—results from, results in—as an umbrella catching the ways one concrete entity can be nothing over and above other entities. The relation thus covers relations like identity, composition, realization, etc. So, for example, the physicalist says, but the emergentist denies, that consciousness results from basic physical entities. And we might agree that a diachronic effect results from its cause when the effect is nothing but a reorganization of the entities present beforehand.)

Emergentism, unlike panpsychism, doesn't issue from a principle of sufficient reason applied across the board. Emergentism tolerates emergent properties that, while triggered by suitable physical entities, don't result from anything, much less their triggering bases. Although an emergent emerges lawfully, the relevant natural law merely records the unbroken pattern between (the coming into existence of the) emergent and (the tokening of the) triggering base. It remains a brute fact that a token base occasions a token emergent. And it remains a brute coincidence that

⁴ For more on emergentism, as well as a critique of it, see Lewtas (2013a).

⁵ Thus we have “mind dust” panpsychism, the kind put forward by James (1890/1950) and Strawson (2006), and debated by Nagel (1979) and Van Cleve (1990). See Lewtas (2013b) and Lewtas (forthcoming b) for more about this view. Some panpsychists—Russellian panpsychists—go on to reduce (so-called) basic physical entities to basic conscious properties (with the result that basic conscious properties aren't instantiated by basic physical entities, but rather supply their underlying categorical grounds). See Lewtas (forthcoming c) for a critique of Russellian panpsychism.

⁶ Other philosophers share this assessment, including most non-physicalists and even some physicalists (namely a posteriori physicalists, Chalmers' Type-B materialists, but not, of course, a priori physicalists, Chalmers' Type-A physicalists; see Chalmers 2003).

every such base occasions, not only an emergent, but also the same kind of emergent.⁷ Still, emergentism, like panpsychism, keeps its metaphysics and epistemology in step. The emergentist naturally enough declines to explain processes (such as property emergence) and patterns (such as the covariation between emergent and triggering base) that happen without reason or cause.⁸ But neither does he posit entities that result from other entities, pursuant to reasons and/or causes, in ways he can't explain. His metaphysic doesn't outrun his epistemology.

This paper argues that physicalist metaphysics does outrun physicalist epistemology, in a way that casts doubt on the rationality of physicalism as an explanatory theory, and therefore also as a metaphysical and ontological theory. It argues, that is, that physicalism is irrational in ways that panpsychism, and even emergentism, aren't. The paper focuses on the much-discussed case of consciousness. Physicalism tells us that consciousness results from physical entities, yet can't explain how. This paper doesn't claim that this makes physicalism wrong (although many of its arguments inevitably tend that way). It maintains, rather, that it makes physicalism irrational. The paper defends this claim by setting forth standards of rationality that, it argues, physicalism must adhere to. Then it shows how physicalism fails to meet them. Sometimes the paper makes general arguments. Sometimes it looks at specific authors as exemplars of physicalist views, then argues, or in some cases suggests, that the point at hand applies more widely. The paper therefore doesn't *prove* its case—whatever, in this context, that might mean. But it hopes to make its case credible enough to give the physicalist pause.

1 Method and Overview

Philosophy has spawned huge literatures about rationality, each featuring a cacophony of competing theories. Nowhere do we find consensus, whether about the nature of rationality, the rationality of agents, the rationality of decisions, the rationality of inferences, the rationality of beliefs, or the rationality of systems of beliefs. This paper would therefore sink into quicksand if it asked, in the abstract, what makes a theory or collective world-view rational and then applied the results to physicalism. The paper therefore pursues a different and three-step approach. First, it shows that the practises and findings of science generate norms of rationality that apply to scientific theories and argumentation. Second, it shows that physicalism necessarily inherits these norms given the particular relation it bears to science. And third, it shows that physicalism inevitably and systematically violates these norms. The paper thus contends that physicalism suffers from incoherence between itself, as a theory, and standards of theory-rationality inseparable from it. The first two

⁷ See Lewtas (2013a) for further defence of this claim.

⁸ The physical base counts as a Human cause of the emergent, because it stands in an appropriate formal relation to it [constant conjunction, counterfactual dependence, perfect correlation, condition necessary and sufficient (or merely sufficient), etc.]. But it doesn't count as a non-Human cause, where we understand non-Human causation as requiring that the natures of the causally related entities result in the causal relation between them. Put crudely, a non-Human cause actively brings its effect about.

steps of the argument unfold in Sects. 2 and 3. Sections 4 through 10 develop the third step. Section 4 introduces the problem consciousness poses for physicalism. Section 5 addresses those physicalists—a priori physicalists—who deny the reality of the explanatory gap (between consciousness and the physical) even as an epistemic matter. Section 6 deals with physicalists who acknowledge the epistemic reality of the gap, but hold out hope that some scientific or conceptual breakthrough will bridge it. Sections 7 through 9 speak to physicalists—the a posteriori physicalists—who accept the gap as an epistemic reality here to stay, but nonetheless deny it metaphysical import. Section 10 asks whether physicalism has led to fruitful research programs. Finally, Sect. 11 ties the paper's arguments together.

2 Physicalism, Science and Rationality

Physicalism does metaphysics by proxy. It delegates the task of exploring the world to the science departments down the hall. It passes the torch on 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. We might say that physicalism consecrates science as philosophically basic and physics as scientifically basic.⁹

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 therefore, and necessarily, inherits much from science. We can divide these into metaphysical and epistemological bequests. They act, for reasons given below, as constraints on the practise of science, and therefore as standards of rationality for science. They become standards of rationality for physicalism given the intimate relation physicalism bears to science. It follows that physicalism undermines itself insofar as it runs afoul of these standards.

The metaphysical bequests are fourfold.

(1) *Metaphysical rationalism* Physicalism adopts science's necessary working assumption that the world, at bottom, makes sense. Underneath the complexity and apparent chaos lies order. Here the physicalist takes the world to obey, or at least to

⁹ Most any physicalist text supports these claims. See Armstrong (1999) and Heil (2003) for particularly clear cases.

cohere with, some (at least moderately) strong version of the principle of sufficient reason (adjusted, perhaps, to make room for such things as quantum indeterminacy). The scientist may, at the end of the day, have to stomach brute entities and/or random or uncaused phenomena. But the more of these he tolerates, the less adequate he will judge his model or theory of the world. Each brute posit will stand out as a loose end—as a hint, maybe, of an underlying order his model or theory has failed to probe, and therefore always (and however unfairly) as a reproach.¹⁰ The scientific outlook, by its nature, thus drives scientists toward a “theory of everything”. In such a theory, non-basic entities sit at the ends of causal chains beginning from, or atop compositional hierarchies resting on, basic entities rooted in mathematical necessity.¹¹ Physicalism commits itself to *metaphysical rationalism* because, as we saw, physicalism necessarily takes the world as science conceives it and gives it.

(2) *Metaphysical reductivism* Physicalism takes on board science’s reductive outlook: the working hypothesis, or empirical inference, that all non-basic entities are nothing over and above basic physical entities.¹² Science presupposes, for instance, that galaxies derive their properties from the basic entities that make them up; it wouldn’t readily entertain the possibility that electrons derive their *basic* properties from the galaxies that contain them. Everyday illustrations of *metaphysical reductivism* abound: scientists investigate complex systems by first tackling their simpler parts (their simpler subsystems); and scientists move toward complex theories about complex phenomena by first framing simpler theories explicitly assuming away many of the complicating factors. *Metaphysical reductivism* gives physicalism a bottom up rather than top down metaphysic.¹³ Note that *metaphysical reductivism* is entailed by *metaphysical rationalism*. This too explains physicalism’s commitment to it.

(3) *Ontological naturalism* Physicalism must limit its substantive ontology (at least its concrete substantive ontology) to the entities acknowledged or required by science. This limits it—as the paper will shortly argue—to (1) entities that causally affect our senses, directly or indirectly, and (2) entities postulated in causal explanation either of entities which affect our senses or of their effects on our senses. Physicalism would lose all meaning if it tolerated non-natural concrete entities.

¹⁰ Witness physicists’ unease with the brute empirical values in the Standard Model. They therefore hope to uncover some rationale for these values. See Greene 2000.

¹¹ Witness efforts by some string theorists to derive basic physical entities from logic and mathematics alone (see Greene 2000).

¹² Levine (2001), 81 considers this one of physicalism’s necessary core commitments. Science may one day abandon reductivism, Levine allows. But then it will falsify today’s physicalism.

¹³ Some philosophers, of otherwise naturalistic bent, endorse monist views according to which all concrete entities derive their ultimate reality, as well as their natures, from the whole, the cosmos, the One, which these thinkers understand as a single all-encompassing entity (see, for example, Schaffer 2010). This paper doesn’t consider such philosophers physicalists. Although they look to science for the detailed natures of the cosmos and its parts, they turn away from presuppositions deeply embedded in science. To this extent, and despite appearances, they count as old-style metaphysicians rather than contemporary physicalists.

(4) *Non-basic mentality* Physicalism draws the empirical inference that mentality (including consciousness) is a non-basic feature of the world. It bases this on the fact that physics has found no unambiguous evidence of basic or bottom-level mentality (including consciousness)—imprecise and controversial talk about a role for consciousness in the collapse of the quantum wave function aside. Science finds clear evidence of mentality (including consciousness) only when special sciences like biology and psychology study the structures and manifestations of complex nervous systems.¹⁴

The epistemological bequests are more diverse. We can group these into primary and secondary bequests. The primary bequests either express general scientific norms or less basic principles that nevertheless play a big role in what follows. There are four of these.

(1) *Epistemological rationalism* Physicalism must, like science, treat the study of nature as a rational activity governed by publicly defensible norms dealing both with collecting and assessing evidence and with deductive, inductive and abductive inferences therefrom—the kinds of norms that underwrite collective, cumulative, reproducible and empirically verifiable epistemic progress. The physicalist can't accept theories grounded, for example, in trance, revelation or mere authority, unless he also justifies these theories in accordance with the usual scientific criteria. For these other methods fall short of the relevant kind of rationality.¹⁵ This last bequest isn't unique to science and its heirs, of course. Science holds forth as but one way of studying nature rationally.¹⁶ But science gives up everything if it abandons its rationalist approach. And so, too, does physicalism, both as a philosophy endorsing metaphysical and scientific realism and, more particularly, as a philosophy piggybacking on science.

(2) *Epistemic warrant* Physicalism necessarily adopts scientific criteria for assessing evidence and theories. Physicalism can't very well defend its metaphysics by enshrining the scientific method as the first and last word, and then hold theories, including philosophical theories, to other, much less lower, epistemic standards.

(3) *Epistemic trust* *Epistemological rationalism* has a general consequence. We must begin with appearances—the way the world seems—and take them at face value unless we have specific, positive and persuasive reasons to believe them misleading. Nothing else makes sense. If we dismissed appearances as such, then the enterprise of knowing and understanding the world couldn't get started. It wouldn't have anything to work with. And once begun, it couldn't home in on a

¹⁴ This contingent empirical result becomes an essential feature of contemporary physicalism. If future science discovered that electrons have consciousness, or that consciousness plays a basic role in the collapse of the quantum wave function, then physicalism, *today's* doctrine, would be falsified. See Levine (2001), Stoljar (2001) and Wilson (2006) for physicalism's necessary insistence on *non-basic mentality*.

¹⁵ The scientist can, of course, be promiscuous within the context of discovery so long as chaste within the context of justification.

¹⁶ Thus Descartes and Leibniz, among many others, approached the world rationally through a priori philosophy as well as science. See Chalmers (1996, 2012) for a generally non-empirical but nevertheless rationalist engagement with the world.

determinate theory. It wouldn't have any way to mark off good theories from bad.¹⁷ This norm applies to any enterprise that accepts the existence of a mind-independent world and explores it rationally.

(4) *Metaphysical/epistemological parity* Physicalism shouldn't, because science doesn't and can't, let its metaphysics run ahead of its epistemology. Science can't overreach because it extracts its metaphysics from, and ties it directly to, its epistemology—as practised through the scientific method. Science dismisses overreaching as unscientific, often using the epithet “mere metaphysics” for so-called scientific theories that lack (adequate) empirical warrant.¹⁸ And physicalism can't overreach because it takes its concrete metaphysics from an *empirical* discipline. Note that *metaphysical/epistemological parity* follows from *epistemic warrant*.

The following five secondary epistemological bequests express either particular consequences of the primary bequests or else norms that play a smaller role in the paper's arguments.

(5) *Parsimony* *Metaphysical/epistemological parity* entails other norms. One such prohibits theoretical posits that outrun data and evidence. *Parsimony* favors hypotheses that posit the fewest (kinds of) entities lacking independent evidence (that is, evidence beyond their explanatory force in the case at hand).¹⁹ Physicalists often support their views about consciousness by applying *parsimony* to mind–body correlations (see, for example, Smart 1959 and Hill 1991).

(6) *Minimal warrant* An even more particular consequence of *metaphysical/epistemological parity* prohibits theoretical posits not backed by any evidence at all. Einstein came to regret postulating the cosmological constant because, at the time he proposed it, no evidence supported it.

(7) *Global cognitive capacity* Physicalism adopts the working assumption, necessary to science, that we humans have what it takes to understand the world. Science can back off locally—in the face of specific, positive and persuasive reasons for a bounded incapacity (thus the ongoing study of our cognitive blind spots and biases)—but can't admit to global incapacity and either justify its practises or put forth its findings as truth. It follows that physicalism can't derive its metaphysic from science and at the same time deny *global cognitive capacity*. Note that *global cognitive capacity* follows, at least loosely, from *epistemological rationalism*.

(8) *Local cognitive capacity* Although science can admit to local cognitive incapacity, its commitment to *global cognitive capacity* obliges it to assume local capacity in the absence of good reasons to the contrary. Such reasons must be

¹⁷ Paul (2012) argues that any metaphysics, and certainly naturalistic metaphysics (including especially physicalism), must accept *epistemic trust*.

¹⁸ Thus Smolin (2007) dismisses string theory as metaphysics rather than science precisely because, in his view, string theory outruns the evidence.

¹⁹ Different philosophers understand *parsimony* differently. This disagreement doesn't matter here.

specific (they must only apply locally), positive (they must rest on empirical evidence and should identify, not only local cognitive failures, but also the precise mechanisms responsible for them), and persuasive (they must show that humans can't overcome or sidestep the relevant incapacities). Again, *local cognitive capacity* follows from *global cognitive capacity*. And again, physicalism's commitment to *global cognitive capacity* in turn commits it to *local cognitive capacity*.

(9) *Fruitfulness* Physicalism, like science, judges research programs by their fruitfulness. Scientists abandon programs, however promising when launched, that go nowhere. *Fruitfulness* follows, albeit loosely, from *epistemological rationalism*.

Science, of course, presupposes and adheres to other norms. Many of these physicalism inherits too. But the thirteen listed above suffice for the arguments made here.

3 The Nature of the Physical

We need to know more about both physicalism and science before we can fully grasp the reach of the above bequests. Physicalism not only claims that high-level entities are (at least token) identical to arrangements of basic physical entities. But more, it must identify them with token basic physical entities *as physics describes them*. Otherwise, as remarked before, 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, for nowhere else do we find the physical as physics describes it.²⁰

What, then, does physical theory say about the world? We note here 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. 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. The scientist then explains his data by postulating properties, or objects with properties, that would cause his data. Hence he hypothesizes new dispositions, or, at most, objects construed solely in terms of new and old dispositions. He 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, and thus scientific, warrant.

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.

²⁰ Lewtas (forthcoming a) defends this understanding of physicalism against alternatives. It also canvasses the problems it faces and the solutions to them.

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 and spatio-temporally located dispositions. Science and physics therefore say nothing about any non-dispositional 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. This means that physicalism, whose substantive metaphysics has warrant only when it marches in step with the substantive conclusions of science, also endorses a purely dispositional, and hence purely causal/functional, world picture.^{22,23}

4 The Problem of Consciousness

Trouble arises when physicalism confronts consciousness. We've seen that physicalism must construe consciousness as a non-basic feature of the world that, metaphysically, (at least token) reduces to basic physical entities (this follows from *non-basic mentality* together with *metaphysical reductivism*). But a physical *understanding* of consciousness proves elusive. This is the notorious explanatory gap, the (at least apparent) fact that we can't see, and can't see that we ever could see, how basic physical entities could result in consciousness.²⁴ Materialists have wrestled with this problem, on and off, at least since the Greek atomists offered their first crude reductions. Contemporary physicalists have attended intensively to it, for over half a century now, without much success—a judgement warranted by the profusion of mutually incompatible physicalist accounts, each endorsed by only a small minority of physicalists, and none of which has significant empirical evidence in its favor.²⁵

This puts physicalists in an uncomfortable position. Imagine a scientist intent on reducing some high-level entity A to certain lower-level B-entities. Try as he might, he can't get the B-entities to account for the properties of A. (The B entities do shed

²¹ Many physicalists have reached this conclusion. See Lewtas (forthcoming a) for references as well as further argument along these lines.

²² Here we understand "causal/functional" as we did above, and thus more widely than in mind-body functionalism.

²³ Some physicalists recognize, within the physical, not only dispositional reality, but also whatever non-dispositional entities, if any, ground dispositional reality (see Maxwell 1978). This paper has given reasons not to classify such views as instances of physicalism. Lewtas (forthcoming a) gives other reasons.

²⁴ See Levine (1983, 1993, 2001) for background on the explanatory gap.

²⁵ See Levine (2001) for an overview and harsh assessment of physicalist efforts tackling consciousness and the explanatory gap.

some light. Our scientist discovers that the key properties of A take somewhat different forms as the B-entities vary. But he sees no way of explaining those key A properties—either their existence or overall nature—in terms of B.) Surely our scientist would abandon his quest, concluding that the A-entity stands over and above the B-entities. His fellow scientists wouldn't take him seriously if he didn't. His field would move on.

The physicalist has three choices here. He can deny the gap, claiming that we can indeed see the the right kind of links between physical and conscious entities. He can hold out hope that tomorrow's better science or some conceptual breakthrough will close the gap. Or he can accept the gap as a sorry fact of life, but of epistemological and not metaphysical import. In this case, and in lieu of explaining consciousness, the physicalist typically explains away his inability to explain consciousness. He traces the gap to some source other than the non-physical nature of consciousness. This paper looks at all three approaches.

5 Denying the Gap

Physicalists who deny the gap deny that consciousness poses an explanandum over and above the causal/functional (understood broadly, as above).²⁶ They thus understand consciousness as a collection of causal activities/dispositions or causally specified roles detectable, in principle, from the third-person perspective. They therefore deny that consciousness has a non-causal/functional qualitative nature uniquely accessible from the first-person perspective.²⁷

Consider Pereboom (2011), for instance. Pereboom suggests that experience is straightforwardly physical. However, its mode of presentation, the way it presents to us, makes experience appear to have properties it in fact 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. Pereboom admits that the following two claims seem intuitive to many:

- (i) Both the physical and introspective modes of presentation represent a phenomenal property as having a specific qualitative nature, and the qualitative nature that the introspective mode of presentation represents the phenomenal property as having is not included in the qualitative nature the physical mode of presentation represents it as having. (13)
- (ii) The introspective mode of presentation *accurately represents* the qualitative nature of the phenomenal property. That is, the introspective mode of presentation represents the phenomenal property as having a specific qualitative nature, and the attribution of this nature to the phenomenal property is correct. (13)

Pereboom accepts the truth of (i), but suggests, given what we know and can rationally believe, that (ii) could turn out false. In other words, it remains

²⁶ These are a priori physicalists, Chalmers' Type-A materialists (see Chalmers 2003).

²⁷ See Lycan (1996) and Tye (1995) for general support of such claims.

an open possibility that introspective representation is inaccurate in the respect that it represents phenomenal properties as having qualitative natures they do not in fact have.... For example, upon seeing the red tomato, Mary [of Jackson’s knowledge argument] introspectively represents the qualitative nature of phenomenal redness in the *what-it-is-like-to-sense-red* way, and it is an open possibility that her representing it in this way attributes to it a qualitative nature that it actually lacks. (14)²⁸

In this case, we can dismiss the explanatory gap as an illusion.

When Mary is in the room, she does not represent phenomenal states in the characteristic introspective way, and she does not appear to have the information required to represent the complete real natures of these phenomenal states by deriving them from what she knows. But it is a serious open possibility that by virtue of her physical knowledge she can nevertheless accurately represent the complete real natures of these states. Phenomenal properties of these states, in particular, might not have the qualitative natures they are introspectively represented as having. Instead, the natures of these properties might accurately be represented by way of Mary’s physical knowledge. If this possibility is actual, then from her physical knowledge, she can derive every truth about the real natures of phenomenal states. (24)

Pereboom confronts an obvious objection.

One may now ask whether the problem for a physicalist explanation of consciousness has merely been shifted from accounting for phenomenal states and their properties to accounting for their introspective phenomenal modes of presentation. (26)

But he believes he has an answer.

In response, there is no less reason to think that the qualitative inaccuracy hypothesis holds for introspective representations of phenomenal modes of presentation than it does for introspective representations of first-order phenomenal states.... The same point might be made for any further iteration of introspective representations of introspective phenomenal modes of presentation. (26)

Pereboom denies that this “engenders an unwelcome infinite regress” (26).

First of all, a mode of presentation can function as a way a subject represents an introspective representation without that subject also representing the mode of presentation itself. [Mary] might, in addition, represent this mode of presentation, but this would be a distinct representation that is not necessitated. If she did represent the mode of presentation, it could be by

²⁸ Pereboom doesn’t insist that his qualitative inaccuracy hypothesis mirrors the truth of things. He claims, more modestly, that it remains an open and rational epistemic possibility. He therefore puts it forth, not as part of a mind–body theory, but as a device to blunt the force of anti-physicalist arguments, like the knowledge and conceivability arguments, which rest, at the end of the day, on the explanatory gap. This paper bypasses this subtlety.

way of a higher order introspective mode of presentation. However, it would again not be necessitated that she also represents this higher order mode of presentation. Furthermore, it's plausible that when someone introspectively represents a sensation of red by way of [an introspective mode of presentation], she will normally, although not necessarily, also represent [that introspective mode of presentation] by a higher order mode of presentation, but only in unusual cases would she introspectively represent that higher order mode of presentation. (26–27)

What's going on here? At first it looks like Pereboom does shift the explanandum—experience as experienced—from the phenomenal property itself to our mode of presentation of it. In other words, it seems that he introduces a new entity, the phenomenal property as it truly exists apart from our experience of it, relabels the original experience an introspective mode of presentation, and then explains the newly introduced and non-experiential entity rather than experience itself (now relabelled an introspective mode of presentation). In this case Pereboom would have changed the subject from the explanandum we care about to something else, something that, at best, bears some relation to the explanandum we care about.

But Pereboom explicitly denies doing this. As he makes clear in the second last passage, he doesn't in fact shift experience out of the phenomenal property and into an introspective mode of presentation of it. He doesn't do this because he also shifts experience out of the introspective mode of presentation and into a higher-order introspective mode of presentation of that. This shifting iterates, he tells us in the last passage, until we get to an introspective mode of presentation not itself introspectively represented. This leaves us to wonder about the source of the qualitative illusions that fill out experience *as experienced* and give rise to the explanatory gap. If the chain of iterations stops at an introspective mode of presentation that we do experience, then, indeed, Pereboom has merely shifted the explanandum in the way he denies. And in this case he has merely relabelled mental entities without offering an explanation of experience as experienced. On the other hand, if the chain of iterations ends with a mode of presentation that we don't experience, an unconscious mode of presentation if you will, then Pereboom identifies an illusion we do experience with a mental state we don't experience. And this move, if we take the identification to heart, does away with the illusion we experience, and with it experience *as experienced*. Pereboom seems to embrace this result: he admits that his hypothesis “is eliminativist about primitive phenomenal properties and those that are accurately introspectively represented, but not about phenomenal properties themselves” (26). So, in the end, Pereboom explains experience by getting rid of it.

How do Pereboom's suggestions fare in light of the standards of rationality sketched in Sect. 2? By eliminating experience as experienced, Pereboom denies manifest data. But it never makes sense to deny data. One can deny interpretations or explanations of data, and thereby, in very limited measure, explain away certain aspects of them. But one can't deny data altogether. Data themselves one starts with. Nobody who seeks a rational understanding of the world can winnow out the data he likes from the data that lead where he doesn't want to go. Certainly no scientist can

turn his back on consistent, insistent and reproducible data, yet hope to model the world truly. Pereboom therefore violates both *epistemological rationalism* (which enjoins the physicalist to employ scientifically acceptable methods of enquiry) and *epistemic trust* (which enjoins the physicalist to start with appearances and to take them seriously unless he has good reasons not to).

Does Pereboom have good reasons to doubt the appearances? First and foremost, does he have *empirical* reasons—that is, scientifically respectable and hence physicalistically acceptable reasons? Pereboom concedes that he doesn't, that “awareness of a discrepancy between the real nature of a phenomenal property and the qualitative nature we introspectively represent it as having seldom, if ever, arise[s]...” (23). He nevertheless tentatively presents two possible examples.²⁹ The first concerns a fraternity hazing, where a blindfolded student, told that a razor will be drawn across his throat, instead has an icicle drawn across his throat. The student briefly feels pain in place of cold. Pereboom suggests that the student “at first misrepresents the qualitative features of the sensation of cold he actually has as qualitative features of pain...” (22). The second example involves Pereboom's daughter, told by her dentist, right before he injects her with Novocain, that he will sprinkle drops of cold water on her gum. Later she tells her dad that she didn't much like the drops of water but that they didn't hurt. Pereboom proposes that “the dentist's suggestion...kept [his daughter] from introspectively representing the qualitative features of the pain state she was actually in as qualitative features of pain...” (23). In each case Pereboom acknowledges that he offers “a controversial analysis but not an implausible one” (22). A critic might maintain that the frat student in fact does experience pain, however briefly, and that Pereboom's daughter in fact never experiences pain, at all. But more to the point, Pereboom's examples, even interpreted his way, give no support to his hypothesis. Pereboom needs examples where a subject *experiences* a phenomenal property without the experience having any of the qualitative features we tend introspectively to represent phenomenal properties as having. But Pereboom offers no such examples. His examples instead feature subjects who experience one kind of “illusory” introspective mode of presentation in place of another. It seems, then, that Pereboom adduces no empirical or scientific evidence in favor of his hypothesis.

What else motivates it? Pereboom frankly acknowledges advancing his hypothesis to reconcile physicalism with the intuitive force of the knowledge and conceivability arguments.³⁰ In other words, the only real “evidence” for his hypothesis is its capacity to save a theory. Here one thinks of Einstein's reasons for postulating a cosmological constant (to save the steady state theory of the universe in face of his own predictions of an expanding universe) and Lorentz's and Fitzgerald's reasons for postulating their transformation equations (to save the aether theory of light in face of the null results of the Michaelson-Morley experiment). In neither case did any empirical evidence directly support the

²⁹ See 22–23.

³⁰ See 3, 7 and 8. Thus “[m]y intention isn't to work out the details of a positive argument for physicalism but rather to assess the prospects for physicalism in the face of the strongest challenges to it...” (7).

hypothesis (in fact, Lorentz's and Fitzgerald's equations went so far as to explain away the lack of empirical support). We can therefore say that Pereboom's hypothesis violates *parsimony* (which enjoins the physicalist to avoid posits that outrun data and evidence) and *minimal warrant* (which enjoins the physicalist to avoid posits not backed by any evidence whatsoever). Since *metaphysical/epistemological parity* (which enjoins the physicalist to keep his metaphysics in step with his epistemology) entails *parsimony* and *minimal warrant*, Pereboom violates *metaphysical/epistemological parity* too. And since *epistemic warrant* (which enjoins the physicalist to assess all theories by scientific criteria) entails *metaphysical/epistemological parity*, Pereboom also violates *epistemic warrant*.

Pereboom bases his hypothesis on the possibility that we systematically fail to interpret our experiences as they really are. He thus proposes a local cognitive incapacity. However, *local cognitive capacity* enjoins the physicalist to reject such proposals barring good reasons to the contrary. Such reasons must be specific (they must only apply locally), positive (they must rest on empirical evidence and should identify, not only local cognitive failures, but also the precise mechanisms responsible for them), and persuasive (they must show that humans can't overcome or sidestep the relevant incapacities). Pereboom offers no reasons to think his hypothesis satisfies these conditions. His hypothesis therefore violates *local cognitive capacity*.

Does the above critique generalize to others who deny the explanatory gap? It certainly applies to those, like Pereboom, who write off the first-person perspective as an illusion (see Rey 1983, for example). But other philosophers deny the gap in different ways. Some purport to explain away the first-person perspective by physically explaining other things, perhaps third-person reports of so-called first-person states (Dennett 1991), perhaps alleged hidden physical properties of phenomenal states (Armstrong 1999). Others dismiss the first-person perspective for its place in an outmoded folk psychology (Churchland 1996). Consider first those who would explain away the first-person perspective by physically explaining something else. Armstrong (1999), for instance, claims that experience has hidden depths inaccessible to experience. We would know experience as physical if only we could access these depths.

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 this view then it becomes much easier to accept that the secondary qualities might have hidden depths to which we cannot penetrate in perception. (129–130).³¹

This makes sense only if experience includes more than we experience. But it doesn't, of course. The thing we wish to understand is experience itself—the

³¹ 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, effectively means experience.

property of experience *as experienced*. If, at the outset, we define experience as experienced to include the unexperienced, then we change the subject, from experience as experienced to an amalgam of experience and something else.³² We thereby risk explaining the unexperienced something else instead of experience itself. And Armstrong does exactly that: later in his (1999) he identifies experience with the hidden physical nature of experience, and elsewhere he identifies experience with physical properties/states which occupy a given functional role (see his 1993). Experience *as experienced* drops from the picture. Dennett (1991) makes similar moves. He changes the subject from experience itself to third-person reports of first-person states, then physically explains those reports. This might undermine the third-person warrant for the first-person perspective. But it leaves untouched the first-person warrant, the phenomenon we set out to explain in the first place. Besides, explaining third-person reports would explain first-person states only if the third-person reports themselves explained first-person states. But if they did, we wouldn't need, in this context, explanations of those third-person reports. We would already have at hand an explanation of experience.

Armstrong and Dennett change the subject and explain something other than they claim to explain. They thereby lose sight of the explanandum of interest and ignore the first-person data for it. They don't deny that data outright, as Pereboom does. But ignoring data comes to the same thing as denying data. Either way the theorist builds his model in defiance of the evidence. It follows that philosophers who change the subject violate both *epistemological rationalism* (which enjoins the physicalist to employ scientifically acceptable methods of enquiry) and *epistemic trust* (which enjoins the physicalist to start with appearances and to take them seriously unless he has good reasons not to).

What about those, like Churchland, who who dismiss the first-person perspective as a posit of a defunct folk psychology? These thinkers treat consciousness as a theoretical object in an explanatory theory: as a hypothesis or explanans explaining something else. But, in this context at least, the first-person perspective—consciousness—doesn't enter the scene as a theoretical posit in explanation of something else. It stands as a basic datum, a ground-level appearance, something each of us encounters through being what we are. It therefore counts as an explanandum in its own right. Now, mistaking explanandum for explanans, in this context, serves as an excuse to sidestep the first-person evidence for experience. Why? Theoretical posits never have direct evidence in their favor. They sit too far from our sensory interface with the world. We endorse them, or not, depending how well they explain the phenomena we introduced them to account for. (When we do acquire direct evidence for a theoretical posit, it thereby loses its status as a theoretical object.) So labelling experience a theoretical posit deftly classes it with entities for which we can't have direct evidence. Philosophers like Churchland thus commit two sins. They confuse explanandum for explanans. And, partly because of that, they turn their backs on basic data. These thinkers therefore violate

³² If we define experience as experienced to include the unexperienced, then we also tumble into incoherence, thereby violating *epistemological rationalism* (which enjoins the physicalist, at the very least, to forswear outright contradiction).

epistemological rationalism (which enjoins the physicalist to employ scientifically acceptable methods of enquiry) and *epistemic trust* (which enjoins the physicalist to start with appearances and to take them seriously unless he has good reason not to).

In sum, this paper provisionally concludes—what seems intuitively plausible in any case—that anyone who denies the explanatory gap must thereby, and without sufficient reason, deny consistent, insistent, reproducible, bedrock data and also a legitimate concrete explanandum. And this means violating key norms of scientific and physicalistic rationality.

6 Hoping to Close the Gap

Physicalists from our second group look to a scientific or conceptual breakthrough to lay bare the links between physical and conscious entities (see Churchland 1988 and, oddly enough, Nagel 1974). But these thinkers lose sight of the fact that the explanatory gap points, not just to our inability today to see such links, but also to our inability to see even the possibility of such links. What grounds this (at least apparent) impossibility? Two facts. (1) Our phenomenal concepts differ from, and apparently refer to different kinds of entities than, the causal/functional concepts that fill out our physicalistic picture of the world. They do this because consciousness presents as having a qualitative nature beyond the causal/functional. Braddon-Mitchell & Jackson (2007) describes this presentational fact:

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 color experience in general, cannot be captured in functional terms, in terms of causal relations. (131)

(2) Our scientific concepts can't reach beyond the causal/functional because they necessarily refer to causally effective entities construed merely as such. This follows from the Sect. 3 arguments about what science does, can do and must do. What possible conceptual or scientific breakthroughs does this leave? Not (1) that consciousness, appearances to the contrary, lacks qualitative nature. As argued in Sect. 5, neither science nor philosophy can reinterpret the appearance of qualitative experience as anything other than qualitative experience without thereby denying, outright, the data given in experience. And no rational enquiry can deny bedrock data. And not (2) that the physical, as science reveals it, includes non-causal/functional qualitative aspects. First, even if such aspects exist as a matter of metaphysical fact, science will never access them, because, for reasons canvassed in Sect. 3, science can only detect, study and postulate causal/functional entities. Second, if future science nevertheless somehow does discover or postulate non-causal/functional qualitative entities, it would thereby falsify today's physicalism. Today's physicalism limits its metaphysic to the kinds of entities discoverable by the *empirical* scientific method. And this method can only detect, study, postulate and acknowledge causes and effects understood as such. Science will therefore take

us beyond the causal/functional only if it shelves the scientific method for, or supplements it with, some other method of enquiry (trance, intuition, revelation, consensus?). But if it does that, it will leave today's physicalism behind. Third, the kinds of revolutionary posits that might underwrite conceptual links between consciousness and the physical seem likely to run afoul of *non-basic mentality*. Think of posits like conscious electrons, conscious categorical grounds underlying physical dispositions, or a basic role for consciousness in the collapse of the quantum wave function. Any of these might close the explanatory gap. But at the cost of disproving physicalism.

In sum, hopes for a scientific or conceptual breakthrough come down to one of three things. (1) Denying the data tomorrow rather than denying them today, even though we can't ever deny data. This violates *epistemic warrant* (which enjoins the physicalist to hold metaphysical doctrines to the same standards to which scientists hold scientific theories) and *epistemic trust* (which enjoins the physicalist to start with appearances and to take them seriously barring good reasons to the contrary). (2) Putting faith in the impossible—in the shape of an identity (token or otherwise) between, on the one hand, something essentially non-causal/functional (because given to us, as data, in a form at least apparently over and above the causal/functional) and, on the other hand, something essentially causal/functional (because construed solely in terms of causal relations). This violates *epistemological rationalism* (which enjoins the physicalist to explore nature using publicly defensible standards of rational enquiry) and *epistemic trust*. (3) Trusting to an unlikely and perhaps impossible scientific revolution that would falsify physicalism even if it happened. This violates *epistemological rationalism* and, in all probability, *non-basic mentality* (which would straightforwardly falsify physicalism).

On top of this, there is something fishy about begging off debate about one's pet metaphysic in hopes of salvation tomorrow. "My theory faces problems; I can't answer them; the criticisms seem ironclad; but I needn't worry because someday someone will find the knockdown reply I don't have now." This dodge works for a while, but wears thin as the wait grows long. Physicalists have reissued their promissory note time and again for over half a century, during which scientific knowledge has doubled roughly every 15 years. Unless the physicalist gives reasons to expect a breakthrough—specific, positive and persuasive reasons more fine-grained than the general march of science—then his trust in the future amounts to a refusal to engage in rational enquiry today. And this, of course, violates *epistemological rationalism* (which enjoins the scientist and physicalist to engage in rational enquiry and nothing but).³³

³³ This isn't merely the counsel of impatience. It doesn't mean, for instance, that we should give up trying to cure cancer because we've fallen short for well-nigh 60 years. We don't have a cure, but we readily see the shape a cure might take. Physicalists have wrestled with consciousness for just as long yet haven't made intelligible even the *possibility* of a physicalistic explanation.

7 Explaining Away the Gap using Metaphysics

Physicalists of the third kind accept the explanatory gap as a principled epistemic barrier but deny it metaphysical significance.³⁴ Consciousness is nothing over and above the physical even though we don't, and won't ever, see how. These physicalists explain away the gap by explaining our inability to see illuminating links between physical and conscious entities. They thereby trace the source of our handicap to something other than the non-physical nature of consciousness. This strategy can't prove that consciousness is physical, of course. But, if successful, it could blunt any motivation to think otherwise. Thus Levine (2001), whom this paper soon examines more closely, thinks the explanatory gap "is a very deep problem" (92), and impossible to bridge given our current and even future conceptual tools (94), with the consequence that there is "an important sense in which we can't understand *how [physicalism] could be true*" (68). And yet he "stop[s] short of endorsing the metaphysical anti-materialist conclusion" (39).

The physicalist has two options here. He can make the metaphysical claim that ontological links hold between consciousness and the physical even though intelligible links don't. We don't see illuminating links because, as a brute matter of metaphysical fact, none exists.³⁵ On the other hand, the physicalist can make an epistemological claim. The ontological links between consciousness and the physical shed just as much explanatory light, in principle, as the ontological links between, say, water and H₂O. However, something about us—some cognitive failing, perhaps—blinds us to it.³⁶ We look at Levine (2001) because he pursues both approaches.

Consider the metaphysical claim first. Here the physicalist postulates brute identities or brute (or strong³⁷) 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. Here the metaphysical impossibility runs deeper than the weak metaphysical impossibility defended by Kripke. Where Kripke merely

³⁴ These are Chalmers' Type-B materialists, much the majority (see Chalmers 2003). See, for example, Hill (1991), Loar (1997) and Levine (2001).

³⁵ Few physicalists embrace this strategy in print, although Chalmers (1996, 2010) claims that many favour it in personal communication and many more are committed to it. We'll see that Levine (2001) toys with it without definitively endorsing or rejecting it.

³⁶ Many physicalists adopt this approach. See, for example, McGinn (1989) and Papineau (2007).

³⁷ Different philosophers use different terms. Chalmers (1996, 2003, 2010) speaks of strong metaphysical necessities. Levine (2001) talks about brute metaphysical necessities (partly because he distinguishes strong metaphysical necessities, which he deems barely acceptable, from brute metaphysical necessities, which, in most passages, he deems unacceptable). See ahead for citations.

constrains the way we can describe a possible world, brute metaphysical necessities constrain the space of possible worlds.

Levine (2001) offers the following characterization of brute identities/necessities and highlights some of their shortcomings.

[I]f we judge a situation metaphysically necessary, then we must have available a representation of that situation relative to which it is conceptually necessary. Another way to put the point is this: I deny the existence of “brute necessities,” metaphysical necessities that transcend logic (where semantic constraints on concepts are understood to be part of logic in the relevant sense). If, for some alleged metaphysically necessary situation, there were no description of it relative to which it was conceptually necessary—no description of it (not just none that we can think of, but none at all) that manifested either formal validity or semantic/conceptual necessity—it would be hard to understand what could ground the metaphysical necessity. At the very least, I don’t see how we could be in a position to judge that the situation in question was metaphysically necessary. (41–42).

The...alternative [the postulation of brute identities/necessities]...just smacks of metaphysical extravagance. It’s certainly not the case that we need appeal to this sort of brute necessity in order to make sense of any of our other modal claims, such as that water necessarily contains hydrogen. Furthermore, I don’t see how we could ever be in a position to actually embrace any particular realization thesis [such as one claiming that a given conscious property is identical to, or is realized by, a given physical property] unless we had the requisite realization theory [such as one showing *how* the physical property is identical to, or realizes, the conscious property].... (45)

Let’s work through the problems Levine notes. First, he questions our epistemic right to embrace any mind–body identity/necessity in the absence of a relevant realization theory. But if we had a realization theory, then we would understand why the identity/necessity holds, in which case it couldn’t be brute. Levine here gets at the utter lack of evidence for brute identities/necessities. Chalmers (2003) sharpens the point by observing that no possible evidence could support them. We can’t know them a priori because, by hypothesis, they constrain the space of worlds independently of logical and hence conceptual possibility/impossibility. And we can’t know them a posteriori because we have a posteriori access only to the actual world, not the space of possible worlds. It seems, then, that brute identities/necessities violate *minimal warrant* (which enjoins the physicalist not to postulate entities unsupported by evidence), *parsimony* (which enjoins the physicalist not to postulate entities that outrun the evidence) and *epistemic warrant* (which enjoins the physicalist to hold his own theoretical posits to the same standards scientists hold theirs).

Second, Levine condemns brute identities/necessities as metaphysically extravagant, reminding us that we don’t need them to explain any other identities (such as the identity of water and H₂O). Here, perhaps, Levine implies that postulating such identities/necessities violates *parsimony* (which enjoins the physicalist not to

postulate entities that outrun the evidence). Perhaps he also worries that the restrictions on possible worlds imposed by brute identities/necessities seem mysteriously non-natural, and hence not the kinds of entities a physicalist ontology can readily accommodate (and thus violations of *ontological naturalism*). After all, what concrete physical entities could ground such modal properties? How could such modal properties figure in, or be required by, concrete causal relations—which relations, Sect. 3 argued, exhaust the nature of the physical world as given by science and as taken up by physicalism?

Third, Levine wonders what could ground brute identities/necessities. We can strengthen this concern by noting that brute identities/necessities seem not only mysterious, but also incoherent. How can property A (a conscious property) and property B (a physical property) stand in a relationship of identity where nothing about A and nothing about B make this intelligible—and yet where, by hypothesis, the 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 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 subset of its properties. Nor do we cognize A/B as A through a subset of its properties plus a 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/necessity be brute—and therefore unintelligible in principle? How can it fail to be straightforwardly intelligible? Here we have violations of *metaphysical rationalism* (which enjoins the physicalist to postulate entities that make sense) and *epistemological rationalism* (which enjoins the physicalist to employ rational methods of enquiry, and which therefore proscribes incoherent posits).

Brute identities/necessities have other shortcomings, ones Levine doesn't mention. First, for the identity/necessity to seem contingent, the “two relata” must appear not to share all their properties. But we must take this appearance at face value barring good reasons—specific, positive and persuasive reasons—for judging it misleading. But we noted above that no direct evidence supports brute identities/necessities. This rules out good reasons for dismissing the appearance that the “two relata” differ. Thus we have violations of *epistemic trust* (which enjoins the physicalist to take appearances seriously) and hence also *epistemological rationalism* (which enjoins the physicalist to employ scientifically acceptable methods and standards of enquiry).

Second, brute identities/necessities hold only at high levels of organizational complexity—almost surely at the level(s) of the conscious properties and the complex physical entities (at least token) identical to them. After all, if the brute identity/necessity resulted from bottom-level entities and/or bottom-level identities/necessities, then we could, in principle, understand it in terms of them. And in this case the higher-level identity/necessity wouldn't, and in principle couldn't, seem

contingent. But this means the identity/necessity can't (at least token) reduce to basic physical entities. And this violates *metaphysical reductivism* (which commits the physicalist to a bottom-up metaphysic) as well as, perhaps, *ontological naturalism* (because an entity that doesn't reduce to basic physical entities has dubious claims to being natural).

Third, brute identities/necessities, postulated in ad hoc defence of a theory but otherwise lacking direct evidence in their favor, stand on a par with Einstein's cosmological constant and the physical processes proposed by Lorentz and Fitzgerald. More, and much like the Lorentz and Fitzgerald equations, brute identities/necessities outrun not only the evidence, but also all possible evidence, since they explain, and thus explain away, their empirical and conceptual invisibility. It follows that brute identities/necessities violate *minimal warrant* (which enjoins the physicalist to postulate only entities supported by evidence), *parsimony* (which enjoins the physicalist not to postulate entities which outrun the evidence) and hence *metaphysical/epistemological parity* (which enjoins the physicalist not to let his metaphysics outstrip his epistemology).

For all this, Levine seems, at the end of the day, to endorse brute identities/necessities. Not because he wants to, or likes them, but because his physicalism leaves him no choice. Levine says the following after making his case for the reality and intractability of the explanatory gap.

First, to avoid commitment to brute metaphysical necessity, I argued that for a situation to be metaphysically impossible, it must be that there is a representation of it relative to which the situation is also conceptually impossible. Second, in order to meet this requirement we would need a realization theory that in the end is based on some identification of the target property so that a redescription of it enables the relevant conceptual necessity to be expressed. (90)

But this is exactly what an unbridgeable explanatory gap rules out. It follows that Levine can't avoid commitment to brute identities/necessities.³⁸ He thus makes all the irrational moves he warns against. And so does any physicalist who discounts the explanatory gap for these kinds of metaphysical reasons. This matters, because brute identities/necessities stand as the only obvious *metaphysical* way to defuse the gap while acknowledging its epistemic reality.

³⁸ The case is a little more nuanced than these arguments suggest. Levine distinguishes between (to use his jargon) brute metaphysical necessities and strong metaphysical necessities (55, 59). Then he argues that strong metaphysical necessities, while carrying heavy theoretical costs, don't offend the canons of good philosophy as badly as do brute metaphysical necessities (58–60). Finally, he denies that his kind of physicalism need subscribe to strong metaphysical necessities anyway (55). For two reasons this paper doesn't accept Levine's pleas of innocence. First, it finds his distinction between strong and brute metaphysical necessities elusive. Second, it judges that Levine commits himself to brute identities/necessities notwithstanding his words to the contrary. Here the paper would repeat the argument made in the text, this time taking into account technicalities in Levine's presentation glossed over above.

8 Explaining Away the Gap using Epistemology I

Physicalists who accept the gap as here to stay also explain it away epistemologically. They claim that the ontological links between consciousness and the physical do shed explanatory light, at least in principle. But some cognitive handicap on our part blinds us to it. These physicalists trace the gap to some form our concepts can't help but take, or to some cognitive faculty responsible for that form. They thereby deny *local cognitive capacity*. This paper looks at several versions of the strategy.³⁹

Perhaps most widely followed is the phenomenal concept strategy, also known as concept or conceptual dualism.⁴⁰ This holds that our concepts cut the world more finely than the properties that fill it out. Thus two disparate concepts, with no a priori connection between them, can home in on the same referent.

Most concept dualists claim that, although our phenomenal concepts do refer to physical entities, they have special features that hide their referents' physical natures from us. It makes sense for the concept dualist to say this. He can't very well allow that our physical concepts fail to represent their referents' natures completely and accurately (subject to gaps or mistakes at the edges, of course). For then he undercuts his warrant for physicalism—which, we have seen, rests significantly on the claim that empirical science yields a complete and accurate account of the world (subject, again, to gaps or mistakes at the edges). Concept dualism therefore pressures the physicalist to blame phenomenal concepts for the explanatory gap.

Concept dualists commonly understand our physical concepts as involving “thick” or descriptive modes of presentation. Thus the concept of an electron attributes the properties physics tells us electrons have. As Levine (2001) puts it, “[most] materialists are willing to grant a cognitively substantive, or “thick” mode of presentation for concepts/terms like “water” – an ascriptive mode that describes water's causal role...” (83).⁴¹ And concept dualists commonly understand our phenomenal concepts as involving “thin” or merely demonstrative modes of presentation. The concept of an experience thus has no more cognitive content than, for example, “that thing over there” or “things like that” taken independently of the entities ostended.

[According to most physicalists] qualitative concepts are essentially “blind” demonstratives. They are pointers we aim at our internal states with very little substantive conception of what sort of thing we're pointing at – demonstrative arrows shot blindly that refer to whatever they hit. (Levine 2001, 84)

Our phenomenal concepts therefore reveal nothing (or at least very little) about the natures their referents have. This explains why brain state 77 seems only contingently related to pain. Our concept “brain state 77” spells out a determinate nature whereas our concept “pain” leaves its referent's nature wholly up in the air.

³⁹ The paper can't look at all of them. New ones pop up all the time. Think of poor Hercules, lopping off the Hydra's heads, only to have two heads spring forth from each severed stump.

⁴⁰ See Alter & Walter (2007) for papers pursuing and criticizing this strategy.

⁴¹ Levine understands modes of presentation as ways entities, or referents, present to us. “One must distinguish between concepts and properties. Concepts are, roughly, modes of presentation of properties (as well as objects)” (46).

Levine, a concept dualist himself, turns this strategy on its head. He considers most concepts, including, especially, our physical concepts, as “non-ascriptive”—as lacking descriptive or other cognitive content and as relating to their referents only by referring to them (rather than, say, by describing them).

[Physical concepts] are “presentationally thin” in the sense that their modes of presentation either contain nothing of cognitive significance beyond the bare representation of the property in question, or contain representations of other properties that are presentationally thin as well. (8)

[So] there is very little, if anything, like conceptual content, or cognitive significance, over and above the actual symbols of the relevant representations and their referents. (53)

Levine understands phenomenal concepts, on the other hand, as “ascriptive”—as having thick modes of presentation that attribute to conscious entities all the richness we experience from the first person perspective.

[O]ur conception, or the mode of presentation of a property like reddishness is substantive and determinate in a way that the modes of presentation of other sorts of properties are not. When I think of what it is to be reddish, the reddishness itself is somehow included in the thought; it’s present to me. This is what I mean by saying it has a “substantive” mode of presentation. In fact, it seems the right way to look at it is that reddishness itself is serving as its own mode of presentation. By saying that the conception is “determinate,” I mean that reddishness presents itself as a specific quality, identifiable in its own right, not merely by its relation to other qualities. (8)

Levine purports to undermine anti-physicalist arguments, not, as most concept dualists do, by pointing to shortcomings in our phenomenal concepts, but instead by trimming back the a priori. “Because I am very sympathetic to the Quinean attack on the analytic/synthetic distinction I would like as much as possible to reduce the a priori to logical form” (42). Given that our concepts “the watery stuff”, “H₂O”, and “brain state 77” lack descriptive content, but our concept “pain” drips with it, the referents of the first three vary across possible worlds (since they can refer to anything) whereas the referent of the fourth stays the same. But then “zombie-H₂O”—H₂O that isn’t watery stuff—is just as conceivable as is an instance of brain state 77 that isn’t a pain (since we can imagine a world where “H₂O” refers to giraffes and “watery stuff” refers to triangles). But we all agree that zombie-H₂O lies beyond the metaphysical pale. It follows, Levine concludes, that “no anti-materialist metaphysical consequences follow from the conceptual possibility of zombies. The materialist has no more to fear from mindless zombies than she has to fear from zombie-H₂O” (54).

Levine nonetheless takes the explanatory gap seriously. He points out that it endures even after we beef up our thin physical concepts with the thick background knowledge we acquire a posteriori about their referents. Thus the referent of our concept “brain state 77”, as understood in the full light of science, fails to explain the referent of our concept “pain”, as understood from the first-person perspective.

In cases...such as “water = H₂O”, while there is no a priori route from the “H₂O”-described facts to the “water”-described facts, still there is the definite sense that when all the chemical facts are in, the whole story has been told.... [N]o sensible question about how H₂O could be water remains.... However, with the proposed identification of reddishness with a physical or functional property...a substantive question does remain.... (91)

These and other instances of concept dualism all make the same kinds of moves. Physicalism comes across data it can't digest. But the concept dualist finds cause to explain away these data as misleading appearances generated by conceptual distortion or blindness. He thus denies *local cognitive capacity* and, to that extent, trims back our overall cognitive rationality. Now, we've seen that the physicalist must presume *local cognitive capacity* unless he has specific, positive and persuasive reasons to the contrary. But the concept dualist has no such reasons. In fact, he advances his conceptual alibi, not because the evidence demands it, but solely to safeguard physicalism. Two facts point this way. (1) We find no consensus among physicalists about whether such incapacities exist. Nor do we find any consensus among concept dualists about what these incapacities are. Each philosopher cooks up his own pet conceptual mechanism, and none of these proposals garners much support from other thinkers.⁴² We see this most clearly with Levine's idiosyncratic concept dualism, which rests on an understanding of physical and phenomenal concepts precisely the opposite of that endorsed by most physicalists. But the other concept dualists bear it out too. Thus Loar (1997) understands phenomenal concepts as recognitional concepts; Papineau (2007) understands them as quotational concepts; Block (2007) understands them as a different and incompatible kind of quotational concept; and Perry (2001) understands them as indexical concepts. (2) Concept dualists consistently adduce philosophical arguments instead of citing scientific facts. Thus Levine attends to the nature and reach of the a priori, the theory of concepts, the epistemic status of both confirmation theory and the theory of reference (62), and the nature of phenomenal concepts as introspectively revealed. The same goes for Loar, Papineau, Block and Perry, although the paper won't support this claim textually. It thus seems that concept dualist hypotheses don't enjoy empirical backing. And yet these hypotheses fall within the domain of cognitive psychology, dealing, as they do, with the contingent nature of our information gathering and processing systems. So the concept dualist defends his philosophical theory against basic data by practising armchair (and largely a priori) science. This violates *epistemic warrant* (which enjoins the physicalist to hold philosophical theories to the same standards to which scientists hold scientific theories). We can compare concept dualism with Freudian psychology, infamous for diagnosing criticisms of it as symptoms of the psychopathological mechanisms postulated by it. Furthermore, because concept dualism postulates cognitive incapacities only to save physicalism, it violates *minimal warrant* (which enjoins the physicalist to postulate only those entities that enjoy evidential support) and *epistemological rationalism* (which enjoins the physicalist

⁴² See the essays in Alter & Walter (2007), the works cited there, and the discussions in Levine (2001) of other concept dualists.

to study nature in accordance with publicly defensible norms about collecting and assessing evidence).

9 Explaining Away the Gap using Epistemology II

Colin McGinn (1989) postulates a cognitive shortcoming—call it M—that keeps us from grasping the straightforwardly physical nature of consciousness. McGinn says we know that a physical psychophysical link obtains—call this P—because we know that physicalism holds true. He then canvasses the only cognitive avenues through which, he claims, we might hope to grasp P—introspection and empirical neuroscience. Introspection won't lead us to P because it doesn't reveal consciousness as depending on neural processes. And empirical neuroscience won't lead us to P because, McGinn argues, scientific concepts are essentially perceptual, perceptual concepts are essentially spatial, and spatial concepts have no grip on consciousness. Hence M.

McGinn makes the same moves as concept dualists—he postulates a contingent cognitive mechanism solely to prove physicalism against basic data, then supports his hypothesis with philosophical argument rather than empirical evidence. But McGinn goes one step further. He does the very thing the theist does when pressed about the problem of evil. He insists that the apparently fatal problem besetting his metaphysic, a problem he can't answer, does have an answer, but one that lies safely outside the reach of human thought. The theist, under duress, assures us that God works in mysterious ways, that our finite minds can't expect to grasp his infinite design. McGinn, a physicalist in a corner, assures us that P would reveal the straightforwardly physical nature of consciousness if only it didn't lie beyond our cognitive competence. But just as the physicalist scorns the theist's dodge, so we might scorn McGinn's. We can't prove him wrong, of course, any more than we can disprove the theist. Who knows what wonders lurk where human minds will never go? But both McGinn and the theist, by seeking refuge in the unknowable and incomprehensible, opt out of rational enquiry. Where the scientist and the philosopher aim to understand the world through rational thought, McGinn and the theist do something else.⁴³ And this something else violates *epistemological rationalism*.

10 Physicalism and Fruitfulness

The physicalist embraces science because it works. And science, for its part, judges scientific theories and research programs by whether and to what extent they work—in a word, by their fruitfulness. Some of these theories and research

⁴³ We can read McGinn more charitably, of course—not as arguing for physicalism, but as exploring what must follow about us from the truth of physicalism coupled with our inability to see its truth. But even this kinder reading reveals metaphysical hubris. Compare McGinn with the creationist, who, confronted with the fossil record, suggests that God brought forth the world six thousand years ago with fossils in place.

programs, like the theory of evolution by natural selection and the research program of evolutionary biology, deliver the goods and flourish. Others, like the caloric theory of heat and the program of phrenology, don't deliver and fall by the wayside.

Is physicalism delivering the goods? Does it make progress on the problems it tackles, resolving one after another, making ever more of the world intelligible in light of its central claims? It seems not. Physicalism may not, in fact, have had any major successes at all. We've seen how it spins its wheels with consciousness. It also spins its wheels in other domains.⁴⁴

This paper obviously can't delve deeply into each research program physicalism spawns. Instead it skims lightly over the surface, summarizing the state of play, noting problems, and drawing a few very tentative conclusions. In a word, this section aims to suggest more than to argue.

Physicalists have struggled to naturalize norms. However, despite much labor, they have yet to field a theory with much plausibility, let alone one that enjoys much support, even among physicalists.⁴⁵ Every effort seems to miss the should-ness, ought-ness, binding-ness—the normative-ness—that sets norms apart.

Physicalists freely admit that indexical facts don't reduce to physical facts.⁴⁶ Most believe that this doesn't threaten physicalism. But physicalism claims that all concrete entities are either physical or nothing over and above the physical. It would seem to follow that (1) indexical facts don't exist (contrary to appearances as well as the beliefs of most physicalists); (2) indexical facts count as facts about abstract entities rather than concrete entities (which makes little sense, especially on a physicalist ontology); or (3) physicalism is false.

Physicalism has difficulty accommodating modality, not least because possibilities and necessities don't readily reduce to actual concrete physical entities.⁴⁷ This has led some physicalists to expunge modality from their ontologies.⁴⁸ It has also motivated most physicalists to seek accounts of causation that shun entities like necessary connections.⁴⁹ But if necessary connections don't obtain, in some form at least, then causal relations can't reduce to their causal relata. They must instead stand over and above the entities that figure in them. Hence the claim, shared by most physicalists, that causal relations and laws of nature hold contingently.⁵⁰ But this seems to turn causal laws, from epistemological artifacts recording the causal capacities of concrete entities, into metaphysical actors with causal clout in their own right. But how can laws push concrete physical entities around? And how can the physicalist naturalize causal laws that don't reduce to—that aren't even grounded in—the concrete physical entities that figure in them? No good answers lie

⁴⁴ See Nagel (2012) for an extended argument to this effect.

⁴⁵ See Miller (2013) for an overview of physicalist meta-ethical thinking.

⁴⁶ See, for example, Perry (1979).

⁴⁷ Lewis (1986) reduces possibilities to concrete actualities, but only at prohibitive ontological cost!

⁴⁸ See, for example, Quine (1961).

⁴⁹ Most physicalists adopt a broadly Human account of causation. See Mackie (1980) and Lewis (1983) for examples.

⁵⁰ Humans about causation are committed to the contingency of both causal relations and laws of nature. See Lewtas (2013a) and Lewtas (forthcoming d) for more about this.

to hand, certainly none that should please the physicalist. The physicalist thus seems faced with a dilemma. If he recognizes modal entities, his physicalism comes under pressure because these entities don't readily reduce to actual physical entities. But if he disavows modal entities, he seems forced to reify causal laws, a move which also strains his physicalism.

Physicalism also has trouble with meaning, whether construed as sense (very roughly, what a thinker or language user grasps) or in terms of reference/intentionality (very roughly, the relation that holds between the thought or expression and the entity the thought or expression is about). Physicalist attempts to naturalize sense run into problems of under-determination of theory by all available and even possible data, where attributions of sense play the role of theory and all possible physical facts play the role of data. Thus Quine and Kripke's Wittgenstein argue from idealized epistemological access to all facts relevant to determinations of meaning, to constitutive meaning skepticism, and finally to meaning nihilism.⁵¹ But we experience linguistic and mental senses in our daily lives. More, our lives, as lived and experienced, wouldn't be possible if these senses didn't exist.⁵² Similar problems arise with intentionality and reference. The physicalist has at hand only one relation naturalistic enough for the task of naturalizing them—causation. But causal relations, understood as actual entities shorn of modal attributes, seem ill-suited to capture the normativity of reference and intentionality. Thus, for example, causal/covariational accounts of intentionality falter in their efforts to accommodate the possibility of error.⁵³ Teleological accounts, such as those offered in Dretske (1981) and Millikan (1984), face similar difficulties. Many physicalists insist that it's early days yet, and that, in time, some causal account will bear fruit.⁵⁴ But others hold out less hope.⁵⁵

None of these laughably brief cartoons carries much weight, of course. But they perhaps suggest that physicalism isn't solving problems, isn't advancing its research programs, and isn't contributing to knowledge the way science and physics do. Physicalism, this paper proposes, has so far proved sterile and shows no sign of getting on track. If this is right, then physicalism violates both *epistemic warrant* (which enjoins the physicalist to hold philosophical theories to the same standards to which scientists hold scientific theories) and *fruitfulness* (which enjoins the physicalist to forsake research programs that don't pan out).

11 Conclusion

It might not have turned out this way. Physicalism isn't irrational from the outset. But the doctrine ran into entities it couldn't make sense of in light of its overarching

⁵¹ See Quine (1960) and Kripke (1982).

⁵² See Lewtas (forthcoming e) for further development of these points.

⁵³ Here we have the problem of ruling out, as illegitimate, incorrect but actual word uses. Suppose someone mistakenly uses "horse" to refer to a cow. How, on causal/covariational accounts, can his word not, on that occasion at least, mean cow? Fodor (1987) discusses these kinds of issues at length.

⁵⁴ Levine (2001) claims this.

⁵⁵ See, for example, Lycan (2008).

metaphysic. This forced physicalists to make moves, and physicalism to take on shapes, that conflict with the metaphysical and epistemological principles the doctrine necessarily inherits from science. At bottom, science and philosophy are ways of engaging rationally with a world itself necessarily taken to be rational (in the sense of being orderly and thus, at least in principle, intelligible). Physicalism, through contingent bad luck, found itself faced with a choice: between giving up its core metaphysic and giving up standards of rationality essential to it. Physicalists opted for the latter. Where science extends human knowledge, and feeds our hopes for understanding more and more of the world, physicalism denies data, attributes brute and mysterious features to the world, cuts back our rational and cognitive capacities, insists on a metaphysic that outruns its epistemology, and embraces dogmatism even unto faith.⁵⁶

References

- Alter T, Walter S (2007) *Phenomenal concepts and phenomenal knowledge*. Oxford University Press, New York
- Armstrong DM (1993) *A materialist theory of the mind*, 2nd edn. Routledge, London
- Armstrong DM (1999) *The mind–body problem*. Westview, Boulder
- Block N (2007) Max Black’s objection to mind–body identity. In: Alter T, Walter S (eds) *Phenomenal concepts and phenomenal knowledge*. Oxford University Press, New York
- Braddon-Mitchell D, Jackson F (2007) *Philosophy of mind and cognition*. Blackwell, Oxford
- Chalmers D (1996) *The conscious mind*. Oxford University Press, New York
- Chalmers D (2003) Consciousness and its place in nature. In: Stich S, Warfield T (eds) *The Blackwell guide to philosophy of mind*. Blackwell, New York
- Chalmers D (2010) *The character of consciousness*. Oxford University Press, New York
- Chalmers D (2012) *Constructing the world*. Oxford University Press, Oxford
- Churchland PS (1988) The significance of neuroscience for philosophy. *Trends Neurosci* 11:304–307
- Churchland PM (1996) *Matter and consciousness*, Revised edn. The MIT Press, Cambridge
- Dennett D (1991) *Consciousness explained*. Little, Brown, Boston
- Dretske F (1981) *Knowledge and the flow of information*. Bradford Books/MIT Press, Cambridge
- Fodor J (1987) *Psychosemantics: the problem of meaning in the philosophy of mind*. MIT Press, Cambridge
- Greene B (2000) *The elegant universe*. Vintage, New York
- Heil J (2003) *From an ontological point of view*. Oxford University Press, New York
- Hill C (1991) *Sensations: a defense of type materialism*. Cambridge University Press, Cambridge
- James W (1890/1950) *Principles of psychology*. Dover Publications, New York
- Kripke S (1982) Wittgenstein on rules and private language. Oxford University Press, Oxford
- Levine J (1983) Materialism and qualia: the explanatory gap. *Pac Philos Q* 64:354–361
- Levine J (1993) On leaving out what it’s like. In: Davies M, Humphreys G (eds) *Consciousness: psychological and philosophical essays*. Blackwell, Oxford, pp 121–136
- Levine J (2001) *Purple haze: the puzzle of consciousness*. Oxford University Press, New York
- Lewis D (1983) New work for a theory of universals. *Austr J Philos* 61:343–377
- Lewis D (1986) *On the plurality of worlds*. Blackwell, Oxford
- Lewtas P (2013a) Emergence and consciousness. *Philos* 88(4):527–553
- Lewtas P (2013b) What it is like to be a quark. *J Conscious Stud* 20(9–10):39–64
- Lewtas P (forthcoming a) *Physicalism and the intrinsic nature of consciousness* unpublished ms
- Lewtas P (forthcoming b) *Building minds: solving the combination problem* unpublished ms

⁵⁶ The author thanks Joe Viste for encouraging him to bring this paper’s arguments together in one place. The paper would never have come about but for him.

- Lewtas P (forthcoming c) Russellian panpsychism: too good to be true? unpublished ms
- Lewtas P (forthcoming d) How to make dualist interactionism work unpublished ms
- Lewtas P (forthcoming e) Consciousness, qualia and cognition: why consciousness includes more than qualia unpublished ms
- Loar B (1997) Phenomenal states. In: Chalmers D (ed) *Philosophy of mind*. Oxford University Press, New York
- Lycan W (1996) *Consciousness and experience*. Bradford Books, Cambridge
- Lycan W (2008) Phenomenal intentionalities. *Am Philos Q* 45(3):233–252
- Mackie J (1980) *The cement of the universe: a study of causation*. Oxford University Press, New York
- Maxwell G (1978) Rigid designators and mind–brain identity. In: Savage C (ed) *Minnesota studies in the philosophy of science*, vol 9. University of Minnesota Press, Minneapolis, p 1978
- McGinn C (1989) Can we solve the mind–body problem? *Mind* 98(891):349–366
- Miller A (2013) *Contemporary metaethics: an introduction*, 2nd edn. Polity Press, Cambridge
- Millikan R (1984) *Language, thought, and other biological categories*. Bradford Books/MIT Press, Cambridge
- Nagel T (1974) What is it like to be a bat? *Philos Rev* 4:435–450
- Nagel, T. (1979). *Panpsychism*. In *Mortal Questions*. New York: Cambridge University Press
- Nagel T (2012) *Mind and cosmos*. Oxford University Press, New York
- Papineau D (2007) Phenomenal concepts and perceptual concepts. In: Alter T, Walter S (eds) *Phenomenal concepts and phenomenal knowledge*. Oxford University Press, New York
- Paul L (2012) Metaphysics as modelling: the handmaiden’s tale. *Philos Stud* 160:1–29
- Pereboom D (2011) *Consciousness and the prospects of physicalism*. Oxford University Press, New York
- Perry J (1979) The problem of the essential indexical. *Nous* 13:3–21
- Perry J (2001) *Knowledge, possibility and consciousness*. MIT Press, Cambridge
- Quine WVO (1960) *Word and object*. MIT Press, Cambridge
- Quine WVO (1961) Reference and modality. In: Linsky L (ed) *Reference and modality*. Oxford University Press, Oxford, p 1971
- Rey G (1983) A reason for doubting the existence of consciousness. In: Davidson R, Shwartz G, Shapiro D (eds) *Consciousness and self-regulation*, vol 3. Plenum Press, New York
- Schaffer J (2010) Monism: the priority of the whole. *Philos Rev* 199(1):31–76
- Smart JJC (1959) Sensations and brain processes. *Philos Rev* 68:141–156
- Smolin L (2007) *The trouble with physics: the rise of string theory, the fall of a science, and what comes next*. Mariner Books, Boston
- Stoljar D (2001) Two conceptions of the physical. *Philos Phenom Res* 62:253–281
- Strawson G (2006) Realistic monism: why physicalism entails panpsychism. In: Strawson G et al (eds) *Consciousness and its place in nature*. Imprint Academic, Exeter
- Tye M (1995) *Ten problems of consciousness: a representational theory of the phenomenal mind*. Bradford Books, Cambridge
- Van Cleve J (1990) Mind: dust or magic? Panpsychism or emergence? *Philos Perspect* 4:215–226
- Wilson J (2006) On characterizing the physical. *Philos Stud* 131:61–99