

8

Why I Am Not a Dualist

Karen Bennett

I am not a dualist. I do not think there are any nonphysical properties, substances, or facts. I think that the entire nature of the world is grounded in—determined or settled by—its fundamental physical nature.

But *why* do I think this? In the bright light of day, I take physicalism to be almost obvious. But in the dark of night, I have to admit to myself that it is not entirely clear why exactly I dislike dualism. Are there good arguments against it? That is, set aside whether there are good objections to the arguments *for* dualism, or against physicalism. Are there good arguments against the view itself?

In what follows, I will take up that question. After more carefully spelling out what I take dualism and physicalism to be, I will suggest that the most frequently heard arguments against dualism are more problematic than we physicalists like to admit. I will then offer a new argument against dualism. In broad strokes, it is this: dualists do not dodge all demands for explanation by denying that consciousness can be explained in physical terms. I will articulate what exactly it is that they must explain, and offer two independent arguments for thinking that they cannot do so. The basic upshot is that moving to dualism because of a perceived explanatory failure of physicalism simply does not help.

1. Dualism

Dualists think that not all the facts are physical facts. They think that there are facts about phenomenal consciousness¹ that cannot be explained in purely physical terms—facts about what it's like to see red, what it's like to

¹ See Block 1995 on the distinction between what he calls 'access consciousness' and 'phenomenal consciousness.' I will usually just say 'consciousness,' but it is the latter I have in mind.

feel sandpaper, what it's like to run ten miles when it's 15° F out, and so on. These phenomenal facts are genuine 'extras,' not fixed, determined, grounded by the physical facts and the physical laws. To use the standard metaphor: even after God settled the physical facts and laws, he had more work to do to put the phenomenal facts in place. Some dualists think that the additional work involves the creation of a special kind of nonphysical substance. More common these days are dualists who think that the additional work merely involves the creation and positioning of special nonphysical properties, and that is the only form of dualism that I will be explicitly concerned with here. The property dualist's claim is that phenomenal properties, or at least protophenomenal properties, are among the basic furniture of the world.

Importantly, however, the property dualist does not propose to ignore the evidence from neuroscience. She does not think that the phenomenal facts float utterly free of the physical facts and laws; she thinks they are connected in important ways. But she thinks these connections are contingent. They are breakable, unlike the connection between, say, being a cat and being a mammal, or that between the existence of some atoms standing in certain complex relations to each other, and the existence of a composite object like a table.² That is how the property dualist maintains a reasonable respect for the physical sciences, while simultaneously claiming that phenomenal properties are genuinely new additions to the world.

Most contemporary property dualists—at any rate, the ones who are my primary target in this chapter—motivate their view by appeal to a family of arguments that are, in the first instance, arguments against physicalism. What I have in mind are the conceivability argument (Descartes 1641; Kirk 1974, 1996; Kripke 1980; Chalmers 1996), the knowledge argument (Nagel 1974; Jackson 1982), and the more general issue that lies in the background of both—the explanatory gap. Both the knowledge argument and the conceivability argument are largely driven by the fact that we don't seem to have any idea how the massively complicated pattern of electrochemical activity in my brain could possibly account for what it's like to see red, or feel sandpaper, etc. As Joseph Levine (2001: 77) puts it, “there seems to be no discernible connection between the physical description and the mental one,

² Of course, not everyone believes in composite objects; some instead endorse what has come to be known as 'compositional nihilism' (including, to varying degrees, van Inwagen 1990; Merricks 2001; and Dorr and Rosen 2002). But most people, including these nihilists, think that the principles that link simples arranged in certain ways to composite objects are necessary, or necessarily false. The exception is Ross Cameron 2007.

and thus no explanation of the latter in terms of the former.” Tell us all the neuroscience you like; it’s still a mystery why *that* is what red looks like. That is why we can apparently conceive of zombies, and why it seems compelling to say that Mary learns something new when she emerges from her black-and-white room. Though the details of the particular arguments differ, the purported upshot is the same—namely, that it is a mistake to think that consciousness can be explained in physical terms.

2. What Is Wrong with Dualism?

As I have already indicated, I am not a dualist. Why am I not a dualist? One way to answer that question would be to lay out what I take to be the problems with the arguments for dualism that I have just sketched. There has been a lot of discussion about where exactly those arguments go awry, and those discussions have yielded fruitful work on the relationship between conceivability and possibility, on the nature of phenomenal concepts, and so forth. However, I want to stick with the question of what is wrong with dualism *itself*. Instead of explaining why I am not convinced by the arguments for dualism, I want to discuss why I am committed to finding fault with them in the first place. This is an important task. I do not want my physicalism to be an article of faith.

Unfortunately, it is closer to an article of faith than most of us are willing to admit. The sad truth is that the extant arguments against dualism are not all that compelling. Here are three.

First, consider a quick appeal to simplicity and Ockham’s razor—that we should make do with as little as possible, and not multiply entities beyond necessity (e.g., Smart 1959). This alone is not going to convince the dualist, who will justifiably claim that she *is* making do with as little as possible. After all, simplicity can only break ties. It can only be wheeled in to decide between two views that both account for all the data—when all else is, as they say, ‘equal.’ But the dualist thinks that all else is decidedly *not* equal. She thinks that physicalism cannot account for all the data, and that making sense of conscious experience requires postulating irreducible phenomenal properties. Legitimately appealing to simplicity here requires having independent reason to think that she is wrong about that. Unfortunately, then, this version of the appeal to simplicity is circular. We need to already have reason to think that the physical facts indeed are sufficient for all the facts before we are entitled to shave with Ockham’s razor (cf., Kim 2005: 125–6).

That said, I do think there is a more sophisticated appeal to simplicity or elegance to be made, and I will do so soon. All I mean to dismiss here is the quick thought that physicalism automatically wins because it has a smaller ontology. While philosophers with a taste for desert landscapes (Quine 1948) will certainly be inclined towards physicalism, that aesthetic preference does not by itself constitute an *argument* for physicalism.

A second argument against dualism might be called the ‘argument from optimistic metainduction.’³ Science has always managed to make do without before. It has never before needed to postulate irreducible nonphysical properties to solve tricky, long-lasting problems, so why here, in this one isolated instance? But even if the dualist grants the premise, this argument is not going to convince her either. She will again say that consciousness is *different*, consciousness is *weird*, and that there is every reason to think that it requires special treatment. It is therefore hard to see how this appeal to the success of science fares much better than the quick appeal to simplicity.

A third argument against dualism is the argument from causal exclusion. If the mental is truly distinct from the physical, how can it have nonoverdeterministic causal power without violating the completeness of physics? Some would say that the nonreductive physicalist has just as much trouble answering this question as the dualist does (e.g., Kim 1989, 1993, 1998; Crane 2001), but they are wrong; nonreductive physicalists have a very plausible solution that dualists cannot properly motivate (see Bennett 2003, 2008). Nonetheless, it is not clear that dualists need to really care about this, because it is not clear that dualists need to think that physics is causally complete.⁴ If they do not, they can duck out of the exclusion problem altogether. We physicalists like the exclusion problem because it gets us from the completeness of physics to physicalism proper; it provides the crucial bridge between the two. Unfortunately, though, it is not entirely obvious why we should think that any dualist would want to get on the bridge in the first place.⁵

³ I owe the name to David Baker; it is, of course, a pun on the “argument from pessimistic metainduction” against scientific realism.

⁴ See Papineau 2001 for an interesting survey of reasons to think that physics *is* causally complete, including a critical discussion of the appeal to conservation of energy.

⁵ I am inclined to think that the argument against *substance* dualism from mental causation is in even worse shape. Princess Elisabeth famously charged that Descartes could not make any sense of “how the mind of a human being can determine the bodily spirit in producing voluntary actions, being only a thinking substance” (letter to Descartes 6/16 May, 1643). But notice that how much force this sort of concern has depends upon what the right account of causation is. The Princess’ objection hits its target if causation requires a connecting process (as in Salmon 1984; Dowe 2000). But it is far from clear that it hits its target if causation merely

Now, I am not saying that none of those three arguments has any force at all. I do think the exclusion problem has *some* force, and that it is important that we be clear that it has more force against a dualist than against a nonreductive physicalist (see Bennett 2008). But all told, we physicalists are perhaps not in as good a position as we like to think. Forget about responding to *objections* to our view; why do we hold it in the first place? What entitles us to our rejection of dualism? Why *am* I not a dualist?

My goal in this chapter is to explore a new answer to that question.⁶ Presumably it will not be knockdown, either, but at the least it will contribute to the cluster of concerns that together constitute the case against dualism. The new objection is basically this: dualists do not excuse themselves from all demand for explanation simply because they deny that consciousness can be *physically* explained. Unfortunately, however, nothing they can offer genuinely addresses this demand in a way that is consistent with their reasons for being a dualist in the first place.

In what follows, I will flesh out the details of that sketch. The core of the complaint is clear enough: dualists owe us an explanation that they cannot provide. This places two tasks before me. First, I need to clarify what it is that dualists need to explain. Second, I need to argue that they cannot satisfactorily do so. In the next section, I address the first task. In the rest of the chapter, I address the second.

3. The Dualist's Project

What is it that dualists need to explain? Care is required here, because of course the whole point of dualism is to claim that certain things are *not* explained—that is, certain things are fundamental. There are at least two kinds of things that dualists will say this about: phenomenal properties and physical-phenomenal correlations. Phenomenal properties are things like

requires counterfactual dependency, Humean 'constant conjunction,' or perhaps even probability-raising. Even the substance dualist can say that pains are reliably followed by stimulus-avoidance behavior, that the behavior counterfactually depends upon the pain, and so forth. If he chooses his theory of causation carefully, he can say that mental-physical causal interaction is entirely unproblematic—while treating it entirely on a par with purely physical causation (see Loeb 1981; Kim 2005; Bennett 2007 for related remarks).

⁶ Or at least *new-ish*. Important precursors include Lycan 1981, Hill 1991, and McLaughlin 2001. See Kim 2005, chapter 5 for discussion.

what it's like to taste coffee, what it's like to feel pain, what it's like to see red. Physical-phenomenal correlations (or psychophysical correlations; I will use these terms interchangeably) are correlations between bodily states and the instantiation of these phenomenal properties. Perhaps such correlations hold across people or even across species; perhaps they only hold within a person over time. But there certainly seem to be some. Scientists do MRI scans, lesion studies, and so forth in order to figure out the 'neural correlates of consciousness.' More prosaically, orange juice tastes one way to people when they first get up, and another rather different way after they brush their teeth. Changes to the chemical environment in people's mouths have a very reliable and replicable effect on the way orange juice tastes, just as the ingestion of chemicals like ibuprofen has a reliable and replicable effect on the way a pain feels. So there are two kinds of things that dualists can say are fundamental: phenomenal properties and psychophysical correlations.

And now we get to what I take to be the crucial question. Just how many of these properties and correlations should the dualist say are fundamental? Is it *some* of them, or is it *all* of them? I claim that a minimally plausible dualism will only say that it is some, not all.

To see why, turn your mind to the view that says that *every single* phenomenal property and *every single* psychophysical correlation is fundamental. This is just an enormous number—a presumably infinite number!—of fundamental posits. The feel of a minor papercut is fundamental experiential property E_1 ; the feel of a slightly worse papercut is fundamental experiential property $E_2 \dots$ and so on and so forth. Similarly for the case of psychophysical correlations. Physical process P is reliably accompanied by a sweet taste. Quite similar physical process P^* is reliably accompanied by a slightly less sweet taste. Once again, the list goes on and on. It would be strange indeed if *each* such correlation had the status of fundamental law.

Certainly, the dualist need not say any such thing. Dualist slogans like 'consciousness is fundamental' can be taken as shorthand for the idea that *some* phenomenal properties, and *some* laws governing the correlations between the physical and the phenomenal, are fundamental.

Compare a purely physical case. Let us suppose that gravity is a fundamental force that figures in fundamental physical laws. This certainly does not entail that *every* fact or generalization about gravity is fundamental! No one thinks that generalizations about the behavior of objects with mass 1 kilogram in the Earth's gravitational field are fundamental, as well as generalizations about the behavior of objects with mass 2 kilograms in the

Earth's gravitational field, as well as generalizations about the behavior of objects with mass 3 kilograms in the Earth's gravitational field... oh, and generalizations about objects with mass 1 kilogram in the *moon's* gravitational field... and so on and so forth. Those principles are instead derived from more fundamental, more general laws.

A minimally plausible dualism will take a similar approach. There is no more reason to think that claims like "physical processes of type *P* are accompanied by flavor sensation of type *F*" are fundamental than there is to think that claims like "a 6 cm³ piece of lead weighs such-and-such on the moon" are. What the dualist should say is that there are some fundamental phenomenal properties and some fundamental laws governing how they are connected to the physical. There is a limited stock of fundamental phenomenal properties and psychophysical correlations or laws, which explain or ground the rest. Some correlations are fundamental law; the others are *derived*. In short, it is implausible for the dualist to fall silent about *all* phenomenal properties and *all* of the connections between the physical and the phenomenal. She will fall silent about some, but she owes us an explanation of others.

This is the more sophisticated appeal to simplicity that I referenced earlier. My claim is not that dualism loses to physicalism just because dualism has a bigger ontology. The claim is rather that a version of dualism loses to a different version of dualism because it has a *vastly* larger ontology that is not in any obvious way necessary.⁷ Dualism is a theory about the world, and should be held to the same standards and aspire to the same theoretical virtues as any other theory about the world. Simplicity and elegance and unification matter. A version of dualism that postulates four fundamental psychophysical laws that explain all the rest of the correlations should, all things being equal, be preferred to a version of dualism that leaves all the correlations brute.

Now, I realize that the dualist can make the same reply as I offered earlier to the quick appeal to simplicity in favor of physicalism. She can dig in her heels, and say that there is no way whatsoever to systematize or explain experience: not only can it not be explained in physical terms, but it cannot be explained in *any* terms. There is just nothing to be said about what

⁷ Note that this appeal to simplicity is consistent with my (2017, chapter 8) and Schaffer's (2015) view that the relevant theoretical virtue is simplicity in the fundamentals.

consciousness is or how it is connected to the physical. Every phenomenal property is fundamental, as is every psychophysical correlation.

Personally, I think such an attitude is bananas, and amounts to giving up completely. But I also recognize that someone who has this attitude will not be moved by that thought. So I will restrict the scope of my conclusion to those dualists who *do* see themselves as having something to say about the so-called “hard problem” of consciousness. For that is what is on the table: dualisms that try to explain consciousness and how it arises from the physical by appeal to a relatively small set of fundamental non-physical properties and laws. This amounts to trying to answer the hard problem, just not in fully physical terms.

This is David Chalmers’ strategy. He agrees with everything I have just said (see particularly Chalmers 1996: 124–9; 1997: 399–400), and does not want to simply fall silent about all the psychophysical correlations. He does want to address the hard problem on which physicalism allegedly founders, and thinks that his dualism can help him answer it. He claims that the impossibility of providing a physical explanation of phenomenal consciousness does not mean that we should give up on the hard problem completely, or conclude that “conscious experience lies outside the domain of scientific theory altogether” (Chalmers 1995: 19). Those are not the right reactions. The right reaction, he says, is to look for a different kind of explanation of consciousness. In particular, the right reaction is to accept that answering the hard problem requires going beyond the physical. It requires an “extra ingredient” (Chalmers 1995: xx)—an ingredient that only a dualist can offer:

Once we accept that materialism is false, it becomes clear that . . . we have to look for a “Y-factor,” something *additional* to the physical facts that will help explain consciousness. We find such a Y-factor in the postulation of irreducible psychophysical laws. (Chalmers 1996: 245)

A physical theory gives a theory of physical processes, and a psychophysical theory tells us how those processes give rise to experience. We know that experience depends on physical processes, but we also know that this dependence cannot be derived from physical laws alone. The new basic principles postulated by a nonreductive theory give us the extra ingredient that we need to build an explanatory bridge. . . . Nothing in this approach contradicts anything in the physical theory; we simply need to add further *bridging* principles to explain how consciousness arises from physical processes. (Chalmers 1995: 20)

The extra explanation-allowing ingredient, then, is a set of fundamental psychophysical laws. These laws are supposed to yield a substantive answer to the hard problem—an answer that no physicalist can provide.⁸

The picture thus far, then, is this. The dualist's project should be—and Chalmers' project indeed is—to provide a distinctively dualist explanation of phenomenal experience and the psychophysical correlations by appeal to a relatively small stock of fundamental phenomenal (or protophenomenal) properties and psychophysical laws. The goal is to use those to systematize, unify, and explain. As Chalmers (1996: 127) says, "the case of physics tells us that fundamental laws are typically simple and elegant; we should expect the same of the fundamental laws in a theory of consciousness." The fundamental psychophysical laws do not themselves link particular patterns of neural activity to easily recognized phenomenal states like a sensation of red, or the smell of dust. They are, instead, simple and general—more like $F=ma$ or $e=mc^2$ —and are used to explain those correlations.

4. My Skepticism: Preliminaries

I am skeptical that the dualist can find anything here that will help. I do not think that she can systematize and unify the correlations without undermining her appeal to the explanatory gap. In Sections 5 and 6, I shall try to make this point in two rather different ways.

My first argument is a methodological one, and has to do with the very idea of the dualist engaging in empirical investigation to continue the search for explanation. The basic thought is that there is a tension in the very notion of a "naturalistic dualism"—not a contradiction, certainly, but an odd

⁸ A quick clarification about how the appeal to fundamental laws or bridge principles is supposed to help. Clearly, Chalmers is not saying that he can get some explanatory mileage out of the claim that each macro-correlation is itself a fundamental law. I have just argued that it would be implausible to claim that each one is fundamental—but it would be far *more* implausible to think that doing so somehow explains why they hold! We cannot explain how and why a physical process-type P is accompanied by searing pain by citing a brute, fundamental law to the effect that P is accompanied by searing pain. That is not an explanation; that is just repeating the explanandum. Quite generally, one cannot *explain* a B*A connection by saying that there is a brute B*A connecting law. The macro-correlations are to be explained, not to do the explaining. So Chalmers' appeal to fundamental laws must involve a certain mismatch between the correlations in the *explanandum*, and the correlations in the laws that constitute the *explanans*. Since a correlation cannot be explained by claiming that it is itself a fundamental law, he needs to postulate some *other* fundamental laws to help do so.

tension that it would be a mistake to ignore. I will claim that it is odd for the dualist to think both that empirical work can help her systematize the teeming swarm of phenomenal properties and psychophysical correlations, and that the explanatory gap poses a real problem for physicalists.

My second argument arises from consideration of the shape and structure of the kinds of laws the dualist would offer. Even if we are careful to take seriously the dualist's *a priori* constraints on what sorts of thing might help, we either make no progress on the hard problem, or else wind up replacing the hard problem with a different problem that is equally hard.

Let me make one preliminary point before getting into either argument. I am about to begin fretting about who can offer what in the attempt to explain consciousness and the psychophysical correlations. It is important to bear in mind that *both* dualists and physicalists might want to postulate 'new'—that is, hitherto unrecognized—entities, properties, or laws to enable them to get somewhere with their explanatory burdens. It would be a mistake to think that only dualists can do so, that postulating anything new counts as the failure of physicalism. Physicalism is not the view that everything logically supervenes on, and can be explained in terms of, the properties, forces, entities, and laws understood by *current* physicists. Physicalism does not assume that current physics is finished. It is, of course, notoriously difficult to define 'physical', and consequently notoriously difficult to decide what sorts of additions to the stock of fundamental laws and entities are physicalistically acceptable.

Here is my guideline for deciding that question: if the addition either *is* conscious experience, or is needed *only* to make sense of conscious experience, physicalism fails. But if the addition is needed to make sense of *both* conscious experience *and* an array of other, paradigmatically physical, phenomena, then physicalism might well be true. For example, it is far from obviously a failure of physicalism to postulate a new fundamental force that explains, say, dark matter, gravity, the surface tension of liquids, *and* consciousness. Consequently, the mere prediction that *current* neuroscience, physics, etc. will not be able to explain consciousness—that something new will need to be discovered—is not itself a commitment to dualism. Commitment to dualism is only incurred when the newly postulated properties or laws have a certain nature. The dualist will take some sort of phenomenal or proto-phenomenal properties to be fundamental, and will postulate new fundamental laws that range over them. The physicalist will do no such thing.

5. The Methodological Argument

Consider the ways in which the dualist is going to search for the new fundamental laws. Here, my focus is simply on the search strategy, rather than any guessed specifics about what the laws might actually *be*. (My second argument will involve suggestions about what rough form the laws would have to take.) The dualist's idea seems to be that we should simply continue doing science, but with the appropriately open mind that comes from giving up the presumption that phenomenality can be explained in physical terms. Chalmers often talks this way. For example, he says that the "liberating force of taking consciousness as fundamental" is that "we no longer need to bash our head against the wall trying to reduce consciousness to something it is not; instead we can engage in the search for a constructive explanatory theory" (Chalmers 1997: 400). The dualist proclaims her dualism and then dives into scientific research to see what turns up.

I frankly do not see what is so liberating about dualism. I cannot see how it makes any difference at all to the course of empirical investigation. And since it does not, taking straightforward empirical investigation to help answer the hard problem undermines the appeal to the conceivability arguments to support dualism in the first place. The latter of those two claims is probably the more controversial one, but let me say something about each in turn.

Both the dualist and the physicalist have a long hard search ahead of them, and the difference between their long hard searches is opaque. The dualist and the physicalist have exactly the same research strategies at their disposal. Both will do a lot of serious neuroscience, and both will pay attention to introspective phenomenology in order to get a better understanding of 'phenomenal space.' Both will run labs, employ postdocs, and apply for NSF funding. Their antecedent commitments will not have any impact on what experiments are available to them, or on what they find. The physicalist research project and the dualist research project do not differ in their methodology or tools, but only in their predicted outcome. That is, the only real dispute is about what they will emerge with at the end of the day. When our intrepid researchers open their laboratory doors several hundred years hence, what will they announce? The physicalist bets that they will announce a solution to the hard problem that only relies upon roughly the sorts of laws, properties or entities that we need to make sense of the straightforwardly physical world. The dualist bets that they will announce a solution that takes consciousness as basic, and invokes new fundamental

psychophysical laws. That is the only difference between them. They disagree about the expected outcome of the very same course of scientific investigation.

This puts the dualist in a rather precarious position. The dualist apparently agrees with the (type A) physicalist⁹ both that we are currently perplexed, and that at the end of science we will not be. But it is odd to claim that no long hard search for a physicalist explanation of consciousness can possibly succeed, yet keep faith in the long hard search for new fundamental laws that will enable a dualist to solve the hard problem. This is particularly odd in light of the fact that it is the *very same long hard search*. Indeed, it is hard to see how this faith in the march of science is consistent with the dualist's appeal to the explanatory gap to support her view.

The dualist is making an *a priori* prediction about the outcome of scientific research. The question is whether she is justified in doing so. She, unsurprisingly, will claim that she is—she will claim that she has *a priori* reason to think that the physicalist research program cannot succeed. That is the point of the conceivability argument, and her appeal to the explanatory gap more generally. But *my* point is that her reliance on those arguments is rendered suspect by her subsequent embrace of empirical investigation. If the dualist thinks that scientific research can uncover hitherto unsuspected truths about the fundamental laws governing psychophysical connections, why should she not also think that it can uncover hitherto unsuspected truths about the physical? That it can generate a deeper understanding of our physicalist tools?

The dualist is endorsing a rather odd pair of propositions here. She is simultaneously insisting that

the fact that we have no idea how to explain consciousness in physical terms is a problem in principle, and there is no point in turning to science to help us,

and that

the fact that we have no idea what the fundamental psychophysical laws are is just temporary, and science will save the day.

⁹ The type A physicalist thinks that any apparent explanatory gap between the physical and the phenomenal is merely a function of our ignorance, and will be closed sometime in the future. See Chalmers 2002: 251–2.

These claims are not straightforwardly incompatible with each other, but there is a real tension between them. Acceptance of the latter should undermine confidence in the former. The more you can see how research in the cognitive sciences can tell us how consciousness arises from the physical, the less secure you should be in your intuition that no purely physicalist story could ever work. All told, then, I suspect that the claim that anything explanatory can be found empirically conflicts with the dualist's reasons for being a dualist in the first place.

6. The Second Argument

My second argument for the claim that dualists cannot make progress on the hard problem is independent of the first. So set aside the methodological concern from the previous section, and suppose there is no tension between the empirical search for systematizing laws and the reliance upon the explanatory gap to dismiss physicalism. What should dualism look like? I begin by claiming that dualism must make some kind of appeal to proto-phenomenal properties—properties that are neither exactly physical nor phenomenal. I then consider two quite different ways of pursuing this appeal, and argue that neither helps at all. On the first, the hard problem simply does not go away; on the second, the hard problem is replaced with a different but equally hard problem.

The first stage of the argument, then, is to argue that the dualist is committed to some sort of protophenomenal properties—or, if that label has unwanted associations, perhaps 'phenomenal minima.'¹⁰ Here is the idea. I have already argued that she should not claim that *every* phenomenal property is a fundamental property. The dualist should not claim that *what it's like to see red* is a fundamental property, and so is *what it's like to see crimson*, and so is *what it's like to see magenta* . . . Slogans like 'consciousness is fundamental' ought not be interpreted as meaning that every single phenomenal property and every single psychophysical correlation is fundamental. Rather, the dualist should pursue the project of explaining some phenomenal properties and psychophysical correlations in terms of a more limited stock of fundamental properties and laws.

¹⁰ I owe the phrase to Ted Sider.

What might this limited stock of fundamentals look like? Well, it's not going to consist of some small set of familiar person-level phenomenal properties. For example, it would just be silly to claim that the only two fundamental phenomenal properties are *what it's like to see red* and *what it's like to taste a particular single malt scotch*. Those two do not exhaust the basic ingredients out of which the rich tapestry of conscious experience is woven! And although those particular examples are perhaps especially arbitrary, I do not see how any other small set of person-level experiences could exhaust the ingredients either. It is not the case that specific bits of person-level conscious experience build all the rest of person-level conscious experience.

The more promising move is for the dualist to claim that there are some unfamiliar, fundamental phenomenal or quasi-phenomenal properties out of which the familiar person-level ones are somehow built. There are common elements that combine and recombine in various ways to generate experience as we know it. Systematizing the relationship between the physical and the phenomenal is a matter of figuring out what those elements are, and what general laws govern their relations both to the physical and to each other. This is what I mean by 'phenomenal minima'. They might be properties of very small entities like carbon atoms, or they might be less-than-fully-phenomenal properties of larger entities like brains or persons. I will often speak in the former way, but I officially leave the matter open. In short: the dualist who shoulders the explanatory task I have set her is committed to postulating some sort of phenomenal minima.

From here, I can see two different ways for the story to unfold. The first is a kind of bridge-principle protophenomenalism. The second is a view that has been called various things, but the label that seems to have stuck is 'Russellian monism' (see Chalmers 1996: 153–5; 2002: 265–7; 2016b; Stoljar 2001). I shall look at each in turn.

On the first picture, the protophenomenal properties occupy in some sense an intervening level between the physical and the phenomenal. They constitute a kind of bridge that connects them. On this view, the new fundamental laws that enable a solution to the hard problem would not be directly between the physical and the phenomenal. They would instead be between the physical and the protophenomenal, and the protophenomenal and the phenomenal. That is, the correlations between the physical and the phenomenal would be given a two-stage explanation that makes reference to an intervening protophenomenal level. First, there are fundamental laws connecting properties like, say, *being a carbon atom* and special

protophenomenal properties. Second, there are special fundamental laws of “mental chemistry” (Nagel 1979: 182; Coleman 2012) that govern the interactions among protophenomenal properties. Put enough carbon atoms together in the right sort of structure, and they will yield a pain.¹¹

This picture has a certain appeal. You almost *can* see how the physical gives rise to consciousness; you almost *can* see how from certain arrangements of *carbon atoms* you get a pain. So, have we an answer to the puzzle? Have we a distinctively dualist explanation of how consciousness arises from the physical?

I say no. And the reason I say no is not, I think, what has come to be known as the ‘combination problem’ for panpsychism, though it is at least a cousin of sorts (recall that this chapter was originally written in 2005–6). The combination problem is originally due to William James, and comes in a variety of guises. One central variety is that “experiences don’t sum,” as Philip Goff (2006) puts it—that little fragments of phenomenality simply cannot by their nature combine into a ‘larger’ whole. Another central variety is that if experience *did* have some sort of quasi-mereological structure, we would notice—but our experiences seem simple and unitary. (See Coleman 2012 for these two versions; see Chalmers 2016a for even more versions. See also Stoljar 2006, Shani 2010, and Montero 2016 for more discussion; this list is far from exhaustive.) My concern is not that there is something special about phenomenal or protophenomenal properties that either prohibits them from combining or somehow would make their combinatorial structure manifest. As I said, I can kind of see the appeal of ‘mental chemistry.’ My concern is rather that I do not see how the hard problem can fail to rearise.

I claim that the bridging version of protophenomenalism faces a dilemma: either a version of the hard problem rearises between the protophenomenal and phenomenal, or else a version of the hard problem rearises between the physical and the protophenomenal. The crucial question is: just how phenomenal are these protophenomenal properties supposed to be?

First, suppose that they are not particularly phenomenal at all. This is a reasonable way to go, at least at first glance. After all, it seems sensible to deny that protophenomenal properties have any of the traditional marks of

¹¹ Note that on this approach, consciousness is not itself fundamental. It cannot be given a constitutive explanation in *physical* terms—so physicalism is false—but it *can* be given a constitutive explanation. Hence my earlier claim that the dualist would only *probably* say that consciousness is fundamental.

the mental. Here are three such marks, which are possessed by standard phenomenal properties like *feeling a searing pain*, or *having a visual impression as of a leafy green tree*. First, there it is something it is like to have them. Second, they are introspectible; we have a certain sort of privileged access to them. Third, that access is arguably incorrigible—although I can be wrong about whether I *do* see a tree, I cannot be wrong about whether I *seem* to see a tree. Dualists like to emphasize all three of these features. They are what make the mental so puzzling. And on this horn of the dilemma, we assume that protophenomenal properties have none of these features. They are not introspectible, incorrigible or not, and there is nothing it is like to have them. But the more we make such apparently reasonable claims, the more the putatively protophenomenal properties look more physical than phenomenal, and the view starts looking more physicalist than dualist. If so, though, we now need a story about how consciousness arises from the *protophenomenal*. Now we need to know how certain kinds of fully phenomenal experience—what it's like to see red, what orange juice tastes like after brushing your teeth—arise from complex arrangements of properties that are not themselves fully phenomenal. The explanatory gap has not been closed; it has just been shunted into the space between the protophenomenal and the phenomenal. The hard problem rears its head there.

We move to the second horn of the dilemma by deciding that that was all a mistake. Perhaps it is wrong to think of protophenomenal properties as being so similar to physical ones; perhaps they really *do* have the marks of the mental. Let us, then, consider the claim that protophenomenal properties *are* introspectible, that carbon atoms have privileged access into their protophenomenal states, and that there is something it is like to be a carbon atom. This move would indeed avoid the concern that we now need an account of how consciousness arises from the protophenomenal. However, it does so at a rather high cost. For one thing, the view is arguably committed to a strange near-panpsychism.¹² Even Thomas Nagel, who is tempted by protophenomenalism of roughly this variety—at least to the extent that it should be “added to the current list of mutually incompatible and hopelessly

¹² The view is not committed to full-blown *panpsychism*, unlike the next version of protophenomenalism to be discussed. For one thing, the view allows that there are fundamental physical particles that are not constituents of conscious beings, and which do not have protophenomenal properties. For another thing, bear in mind that the fundamental protophenomenal properties might be possessed only by large and complex physical systems—brains, for example. This, of course, makes them rather different than other fundamental properties, but that is only to be expected.

unacceptable solutions to the mind-body problem” (1979: 193)—resists the idea that “the components out of which a point of view is constructed would . . . themselves have to have a point of view” (194). However, panpsychism is not my real complaint at the moment (I shall say more about it shortly). The important point for the moment is that this view, like the alternative, simply pushes the hard problem elsewhere. **If protophenomenal properties are so like phenomenal ones, well, then now we need a story about how the protophenomenal arises from the physical. We have lost out on the project of explaining personal-level fully phenomenal properties and correlations with less than fully phenomenal ones.**

Either way, then, the protophenomenalist has failed to address the hard problem. The more similar the protophenomenal properties are to phenomenal ones, the less headway can be made on the project of systematizing the correlations. And the more removed the protophenomenal properties are from phenomenal ones, the less point there is to postulating them at all. We still cannot see how human experience—genuine, full-blown consciousness—arises from complicated relations among such fragmentary shadows of phenomenality.

Indeed, there is a case to be made that this bridging version of protophenomenalism slides into a regress. To see it, consider the version that says that protophenomenal properties are quite different from ordinary phenomenal properties—the version on which they do not bear the marks of the mental. (A similar issue arises for the other version.) In order to bridge the gap between the protophenomenal and the truly phenomenal, maybe we should posit a *fourth* kind of property—protoprotophenomenal properties, or, better, protophenophenomenal properties. These occupy the intervening level between the protophenomenal and the phenomenal, and their connections between the two are governed by a limited set of fundamental protophenomenal–protophenophenomenal laws, and fundamental protophenophenomenal–phenomenal laws fundamental. Lather, rinse, repeat. I am not fully convinced this regress argument works, but I nonetheless place it on the table for inspection. My main claim is the one from the previous paragraph: for bridging protophenomenalism, the hard problem rearises either between the physical and the protophenomenal, or between the protophenomenal and the phenomenal.

It is time to move away from the bridging version of protophenomenalism, and on to the second version: Russellian monism. On this view, the phenomenal minima do not occupy an intervening level between the physical and the fully phenomenal, but rather occupy physical properties

themselves. Metaphorically speaking, they form the *core* of physical properties. More carefully, the idea is that there is independent motivation for the view that physical properties and entities can be characterized only relationally, by their causal-dispositional roles (Russell 1927). If such a view is correct, there is a pressing question about what intrinsic properties fill these causal-dispositional roles. One answer to this question is designed to also address the hard problem. If protophenomenal properties fill the causal-dispositional roles, two problems are solved at once. There are various ways to flesh out the details, but all that really matters for my purposes is the view's central claim—the world is qualitative *all the way down*.

Russellian monism avoids the dilemma that faced the bridge version of protophenomenalism; it does not simply push the hard problem elsewhere. Two features allow it to do this. First, the very nature of physical properties and entities is protophenomenal. Physical properties are relational, dispositional, “structural/dynamic” (Chalmers 2002: 265); intrinsic protophenomenal properties underlie them. This means that there is no gap between the physical and the protophenomenal in the first place, and Russellian monism dodges the second horn of the dilemma. Second, as long as the protophenomenal cores are taken to have the marks of the mental, or at least some approximation thereof, there may not be any particularly difficult question of how full-blooded phenomenal properties arise from them. So Russellian monism can dodge the first horn of the dilemma as well.

It is worth taking a moment to emphasize that the Russellian monist *must* claim that the protophenomenal properties are recognizably phenomenal. One reason is that just mentioned—the view would otherwise be impaled on the first horn of the dilemma. But the Russellian monist has a further reason, one that does not quite apply to the ‘bridge’ version of protophenomenalism. This further reason is that there would otherwise be very little reason not to count the view as a form of physicalism. After all, the view is that there are rock-bottom features of the world that account for the charge of electrons, the behavior of gases, the hardness of diamonds . . . and consciousness. This is straightforwardly physicalist if those rock-bottom features are non-qualitative.

Recall my earlier remark that a good guideline for deciding whether or not an addition to our ontology counts as physical is the range of phenomena for which it accounts. If the additional feature either is consciousness, or explains nothing but consciousness, then that is probably sufficient for it not to be physical; if it explains clearly physical phenomena as well, then that is at least a *prima facie* reason to say that it is indeed physical. The Russellian

monist's protophenomenal properties meet that *prima facie* guideline for counting as physical—they ground every physical property. However, if they also bear the marks of the mental, they meet the sufficient condition for *not* counting as physical. If the Russellian monist does not want to be a physicalist, then, he must say that protophenomenal properties are themselves recognizably phenomenal. (See Chalmers 2016b for discussion of whether Russellian monism counts as physicalist or not.)

As I have already suggested, this is not a particularly natural view. It is rather odd to claim that there is something it is like to be a carbon atom. However, it is hard to see how to do more than trade intuitions about this point. So let us set it aside, and suppose that there is, indeed, something it is like to be a carbon atom. I still do not like Russellian monism any more than I like the other version of protophenomenalism. It may solve the official hard problem, but only by generating a new problem that is just as hard.¹³

Start by noticing that the Russellian monist is committed to the following claims. There is no in principle difference between me and a carbon atom, or me and my socks. There are differences in organization, and complexity, and the like, but that is all. These are differences in degree, not kind; there is no unbridgeable chasm between me and my socks. But those, of course, are claims that any *physicalist* will endorse as well. The Russellian monist says that the world is mental all the way down. The physicalist says that it is physical all the way up. Both are forms of monism; both assimilate one of the allegedly different categories to the other.

Of course, the Russellian version of monism does get to say that there is no issue about how the qualitative arises from the nonqualitative, because the world is qualitative all the way down. But now the opposite problem arises! What shall he say about how the *nonqualitative* arises from the qualitative? The physicalist says it is nonqualitative all the way up, and faces a question of how the qualitative gets into the picture. The Russellian monist says it is qualitative all the way down, and thus faces a question of how the nonqualitative gets into the picture. The difference between the physicalist and the Russellian monist is a reversal about what explains what: the physicalist wants to explain the mental in physical terms, and the

¹³ It is tempting, but I think ill-advised, to try to raise another objection here—namely, that the Russellian monist has to deny that zombie worlds are conceivable, and thus has no reason not to be a physicalist. For a nice discussion of the Russellian monist's options, see Chalmers 2002: 266; 2016b: 28–9.

Russellian monist wants to explain the physical in mental terms. Yet the latter is just as tricky as the former, and to my knowledge no Russellian monist has ever even tried to say anything to alleviate the mystery.

After all, the Russellian monist not only claims that there is something it is like to be a carbon atom, but also that its phenomenal character is what *makes it be a carbon atom in the first place*. It is its intrinsic phenomenal or protophenomenal nature that is responsible for all of its causal powers, and that plays the dispositional role associated with being a carbon atom. Its intrinsic (proto)phenomenal nature grounds its disposition to bond in certain ways with hydrogen atoms and so forth, in the same way that possession of a particular crystalline structure grounds a glass's disposition to break if dropped (see, e.g., Chalmers 2002: 265). I have absolutely no idea how this is supposed to work, or why it is supposed to sound plausible, other than the fact that it would be convenient if it were true. So Russellian monism faces an inversion of the standard explanatory gap: I cannot see how to get the nonqualitative out of the qualitative in the way that Russellian monism requires. How can phenomenality be the right sort of thing to explain how negative charge works, the various ways that carbon atoms bond with oxygen, and the like? How would consciousness ground those causal powers?

So much, then, for Russellian monism. Postulating an intervening level of protophenomenal properties, à la the bridging version of the view, relocates the hard problem. Postulating an underlying level of protophenomenal properties, à la Russellian monism, turns the hard problem on its head for no good reason.

7. The Final Moral

Here, again, is the overall picture. I have argued that dualism does not avoid all explanatory burdens. It is very implausible for the dualist to go no further than postulating an enormous proliferation of fundamental properties and unsystematized fundamental psychophysical laws. Minimally plausible, scientifically respectable dualism will instead posit a limited number of fundamental phenomenal or protophenomenal properties, and a limited number of fundamental principles governing physical-phenomenal connections. Dualism will explain some aspects of conscious experience in terms of others, and will explain some particular physical-phenomenal correlations in terms of a few fundamental laws.

However, I have provided two more-or-less independent arguments against the claim that she can make any real progress here. First, I argued that there is a real tension between the dualist's faith in the empirical search for such laws and her deep skepticism about the physicalist's search for an explanation of consciousness in physical terms. The former should undermine the latter. Second, I looked at some possible forms her fundamental laws might take. Even though 'phenomenal minima' like protophenomenal properties might appear tailor-made for closing the explanatory gap, they do no such thing. The bridging version of the view cannot in principle give the dualist any explanatory purchase, and the Russellian monist version dismisses the hard problem at the expense of raising a new one. All told, then, matters look fairly bleak. Consciousness looks at least as mysterious to the dualist as to the physicalist; explanatory gaps remain.

Here is the dialectic as I see it. The dualist challenges physicalism by appeal to arguments like the zombie argument and the Mary argument. In the face of these challenges, the physicalist has a choice about how to respond: he can either cave or resist. And there is a real question, I think, about which of those responses is the correct one. This is an instance of a more general meta-philosophical issue. When should we stick to our guns and defend a view against an objection that is not obviously and straightforwardly fallacious? How stubborn should we be? I do not know how to answer that meta-philosophical question in full generality, or even whether there could be a general answer. But this chapter is intended as an argument for stubbornness in this particular case. Physicalists face challenges from the explanatory gap, yes, but dualists face their own versions. Since the same problems just get pushed elsewhere, we physicalists have motivation to resist. We should hold fast, and endorse one of the many ways of responding to the dualist's arguments.

So, admittedly, I do not know the physicalist solution to the hard problem. I am not sure how the physical facts explain the phenomenal facts. But I cannot see how denying that the physical facts explain the phenomenal facts makes life any easier. *Both* physicalists and dualists face versions of the explanatory gap, and retreating to dualism simply raises further questions that are just as hard as the ones physicalists face. *Dualism simply does not help*. It offers no advantage over physicalism.¹⁴

¹⁴ This chapter has an unusual history. I initially wrote it in roughly 2005–6, and then I shelved it for *fourteen years*. There were reasons for that, though not very good reasons. I have revised it for clarity, and to make contact with some literature that has appeared since I first

References

- Bennett, Karen. 2003. Why the exclusion problem seems intractable, and how, just maybe, to tract it. *Noûs* 37: 471–97.
- Bennett, Karen. 2007. Mental causation. *Philosophy Compass* 2: 316–37.
- Bennett, Karen. 2008. Exclusion again. In Jakob Hohwy and Jesper Kallestrup, eds., *Being Reduced*. Oxford: Oxford University Press, 280–305.
- Block, Ned. 1995. On a confusion about a function of consciousness. *Behavioral and Brain Sciences* 18: 227–47.
- Cameron, Ross. 2007. The contingency of composition. *Philosophical Studies* 16: 99–121.
- Chalmers, David. 1995. Facing up to the problem of consciousness. Reprinted (1997) in Jonathan Shear, ed., *Explaining Consciousness: The Hard Problem*. Cambridge, MA: MIT Press, 9–32.
- Chalmers, David. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Chalmers, David. 1997. Moving forward on the problem of consciousness. In Jonathan Shear, ed., *Explaining Consciousness: The Hard Problem*. Cambridge, MA: MIT Press, 379–422.
- Chalmers, David. 2002. Consciousness and its place in nature. Reprinted (2002) in David Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press, 247–72.
- Chalmers, David. 2016a. The combination problem for panpsychism. In Godehard Bruntrup and Ludwig Jaskolla, eds., *Panpsychism: Contemporary Perspectives*. Oxford: Oxford University Press, 179–212.
- Chalmers, David. 2016b. Panpsychism and panprotopsychism. In Godehard Bruntrup and Ludwig Jaskolla, eds., *Panpsychism: Contemporary Perspectives*. Oxford: Oxford University Press, 19–47.

wrote it. But I cannot claim that I have managed to bring in *all* of the relevant more recent literature, and I apologize to anyone whose work I should cite but do not.

In its earlier incarnation(s), it was presented at the NYU Mind and Language seminar, the Australasian Association of Philosophy conference, the Australian National University, Harvard University, Brown University, Columbia University, the University of Vermont, the University of California Davis, and the Metaphysics of Science workshop at Birmingham University. Thanks to everyone, particularly to the Corridor reading group, David Chalmers, Tyler Doggett, Derk Pereboom, Daniel Stoljar, and Ted Sider for extensive discussion. Thanks also to Selim Berker, Ned Block, Hartry Field, Kit Fine, Michael Glanzberg, Paul Griffiths, Thomas Nagel, Nick Shea, and Susanna Siegel for helpful comments. I'd particularly like to thank the unnamed person who got me started thinking about this chapter by remarking, "But I'm a dualist! I can say anything I want!"

- Coleman, Sam. 2012. Mental chemistry: combination for dualists. *Dialectica* 66: 137–66.
- Crane, Tim. 2001. The significance of emergence. In Carl Gillett and Barry Loewer, eds., *Physicalism and its Discontents*. Cambridge: Cambridge University Press, 207–24.
- Descartes, René. 1641. *Meditations on First Philosophy*. Translated by Donald Cress (1993). Indianapolis: Hackett.
- Dorr, Cian and Rosen, Gideon. 2002. Composition as a fiction. In Richard M. Gale, ed., *The Blackwell Guide to Metaphysics*. Oxford: Blackwell, 151–74.
- Dowe, Phil. 2000. *Physical Causation*. Cambridge: Cambridge University Press.
- Goff, Phillip. 2006. Experiences don't sum. *Journal of Consciousness Studies* 13.10–11: 53–61.
- Hill, Christopher. 1991. *Sensations: A Defense of Type Materialism*. Cambridge: Cambridge University Press.
- Jackson, Frank. 1982. Epiphenomenal qualia. *The Philosophical Quarterly* 32: 127–36.
- Kim, Jaegwon. 1989. The myth of nonreductive physicalism. Reprinted (1993) in *Supervenience and Mind*. Cambridge: Cambridge University Press, 265–84.
- Kim, Jaegwon. 1993. The nonreductivist's troubles with mental causation. Reprinted (1993) in *Supervenience and Mind*. Cambridge: Cambridge University Press, 336–57.
- Kim, Jaegwon. 1998. *Mind in a Physical World*. Cambridge, MA: Bradford.
- Kim, Jaegwon. 2005. *Physicalism, or Something Near Enough*. Princeton: Princeton University Press.
- Kirk, Robert. 1974. Zombies vs. materialists. *Aristotelian Society Proceedings Supplement* 48: 135–52.
- Kirk, Robert. 1996. Strict implication, supervenience, and physicalism. *Australasian Journal of Philosophy* 74: 244–56.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Levine, Joseph. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.
- Loeb, Louis. 1981. *From Descartes to Hume*. Ithaca: Cornell University Press.
- Lycan, William. 1981. Psychological laws. *Philosophical Topics* 12: 9–38.
- McLaughlin, Brian. 2001. In defense of new wave materialism: a response to Horgan and Tienson. In Carl Gillett and Barry Loewer, eds., *Physicalism and its Discontents*. Cambridge: Cambridge University Press, 208–21.

- Merricks, Trenton. 2001. *Objects and Persons*. Oxford: Oxford University Press.
- Montero, Barbara. 2016. What combination problem? In Godehard Bruntrup and Ludwig Jaskolla, eds., *Panpsychism: Contemporary Perspectives*. Oxford: Oxford University Press, 215–29.
- Nagel, Thomas. 1974. What is it like to be a bat? *Philosophical Review* 83: 435–50.
- Nagel, Thomas. 1979. Panpsychism. In *Mortal Questions*. Cambridge: Cambridge University Press, 181–95.
- Papineau, David. 2001. The rise of physicalism. In Carl Gillett and Barry Loewer, eds., *Physicalism and its Discontents*. Cambridge: Cambridge University Press, 3–36.
- Quine, W.V.O. 1948. On what there is. *Review of Metaphysics* 2: 21–38.
- Russell, Bertrand. 1927. *The Analysis of Matter*. London: Kegan Paul.
- Salmon, Wesley. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Schaffer, Jonathan. 2015. What not to multiply without necessity. *The Australasian Journal of Philosophy* 93: 644–64.
- Shani, Itay. 2010. Mind stuffed with red herrings: why William James' Critique of the mind-stuff theory does not substantiate a combination problem for panpsychism. *Acta Analytica* 25: 413–34.
- Smart, J.J.C. 1959. Sensations and brain processes. Reprinted (2002) in David Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press, 60–8.
- Stoljar, Daniel. 2001. Two conceptions of the physical. *Philosophy and Phenomenological Research* 62: 253–81.
- Stoljar, Daniel. 2006. *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. Oxford: Oxford University Press.
- van Inwagen, Peter. 1990. *Material Beings*. Ithaca: Cornell University Press.