

8

Qualia, Qualities, and Our Conception of the Physical World

HOWARD ROBINSON

1. THE REAL POWER OF THE KNOWLEDGE ARGUMENT

The Initial Predicament

The dialectical situation in which the knowledge argument (KA) for property dualism is usually taken to be located is the following.¹ It is taken as agreed that physicalism gives an adequate account of non-conscious reality, and that this part of reality constitutes almost 100 percent of the universe. Despite this overwhelming success, however, the physicalist account struggles to accommodate certain features of mental life, namely the ‘what it is like’ or qualia of certain conscious states. These qualia constitute the qualitative nature of sensations and probably of secondary qualities, but have nothing to do with our robust conception of the physical as it applies to the vast mindless

tracts of reality. These awkward entities constitute what Chalmers called “the hard problem” for physicalism (Chalmers 2003). But the fact that they also constitute such a tiny part of the world is implicitly understood as being a strong *prima facie* reason for thinking that there must be some way of reconciling their apparent existence with the otherwise triumphant and clearly adequate physicalist account of the world: if it were not for the qualia that occur in a few corners of reality, the adequacy of physicalism would not in any way be in dispute.

I think that this interpretation of the situation constitutes a radical misunderstanding of and understatement of the problem that faces physicalism and the role that the knowledge argument plays in bringing out that problem: the dialectic is quite different from the way it is represented in the previous paragraph. To see why and how this is so, one must direct attention at our conception of matter and the physical, rather than at our concept of mind. Science, whether of the macroscopic or the microscopic, is very largely concerned with measurement and quantification and with the expression of its findings in mathematics, as far as is possible. But the resultant abstract—we might call it Platonistic—conception of the physical cannot, we think, wholly capture our concept of the physical, especially as it is conceived in our naïve or commonsensical conception of the world. Taken in this abstract form, the concept of the physical is insufficiently concrete. But what concretizes it is the addition of qualities—essentially sensible qualities—that figure so importantly in our naïve or commonsensical conception of the world. These are essential to our ability to ‘cash’ or ‘model’ or ‘interpret’ the abstract, mathematical conception.

Physicalism’s real predicament, as has been brought out by the KA, can be represented in two propositions.

- (1) Standard physicalism cannot capture the *qualitative* nature or aspect or reality.
- (2) The qualitative is an essential feature of any conception of the physical that goes beyond the purely abstract and mathematically expressed.

These two together entail

(3) Standard physicalism cannot capture any conception of the physical that goes beyond the purely abstract or mathematically expressed.

This is, of course, a much stronger conclusion than that which the normal understanding of the ‘hard problem’ attributes to the knowledge argument, namely

(4) Standard physicalism cannot capture the qualitative nature of certain mental states.

On my reinterpretation of the situation, what the knowledge argument really shows is (1). I take (2) to be independently plausible, possibly analytically true and probably largely uncontested. Propositions (1) and (2) together show that standard physicalism is not merely incomplete, failing to cope with consciousness, but something more like incoherent, because it cannot give a coherent account of the physical itself.

In the next two sections I shall do the following. First, I shall show that the knowledge argument, if sound, proves that physicalism cannot capture the qualitative at all: that is, I shall try to prove (1). Second, I will argue that this does not merely strengthen the knowledge argument’s conclusion, but also undercuts all known attempts to refute the argument, for they all rest on the assumption that the physicalist’s conception of the purely physical is itself unproblematic; that is, they rest on the assumption that the physicalist’s conception of the physical would not be inadequate if it were not for the need to explain consciousness.

Extending the Scope of the Knowledge Argument

The knowledge argument as traditionally stated appears to concern only the nature of mental states. This appearance is founded on two factors. First, the argument concerns ‘what it is like’ to have certain experiences, and this expression clearly names something subjective. This alone is not enough to confine the topic to the mental world. If one is talking about *what it is like* to feel pain or jealousy, then this might seem to be purely internal, but *what it is like* to see color or to

hear sound is directly connected with our notion of *what color is like* or *what sound is like*, as those things are, or as we naïvely conceive them to be, in the external world. There is an irony here. It is a feature of the accounts many physicalists present of experience that it is transparent. This means that the only feature that characterizes the experience *qua* experience is the apparent presence of some objective or external property. So *what it is like* to experience red or C-sharp is no different from *what red is like* or *what C-sharp is like*. This transparency alone is not enough, however, to show that the knowledge argument concerns our conception of the external or physical world, because the qualities that are invoked in the argument are secondary qualities and, at least since Locke, it is standard within the scientific form of physical realism to treat secondary qualities, insofar as they are not just powers or dispositions but monadic qualities, as subjective. This leaves the physical world untouched, for that is characterized wholly by primary qualities.

Once one has reached this point, it ought to be becoming clear why the argument does not concern secondary qualities alone. ‘Red’ is defined in terms of what it is like to perceive the color red; ‘square’ is not defined in terms of what it is like to perceive a square. But someone’s conception of ‘square’ is not independent of what it is like for him to perceive (see, touch, or whatever else possible sense) square things, for if it were it would be wholly axiomatic and mathematical. Using Sellars’s convenient terminology, we can say our ‘manifest image’ of the world is a projection of what it is like to perceive it, in respect of both primary and secondary qualities. The secondary qualities are attached to a particular form of experience, the primary are not. But without any experience—in our case, visual or tactile or both—there would be no conception of spatial properties beyond the wholly mathematical.

What the knowledge argument really brings out is that only experience of the appropriate kind can reveal the qualitative, as opposed to purely formal and structural, features of the world. What the standard modern physicalist fails to notice is that the kind of thing that Jackson’s Mary (Jackson 1982; Ludlow et al. 2004) did not know, generalized from color vision to all the other sensible qualities, is essential to any contentful conception of the world, and hence that

physicalism without it would lack any empirical content. The generalization of the knowledge argument can be expressed as follows.

Take any property *P* which is a quality or has a qualitative aspect; then of any subject *S* who has no experiential grasp on that qualitative aspect but otherwise has full knowledge of all matters relating to *P*; it will be true that *S* lacks knowledge of *P*'s qualitative aspect.

I call this the *generalization* of the argument, but one might, with equal justice, say that it is the principle underlying the argument. The thought experiment merely makes its truth clear.

It is vital to appreciate that this rationale applies to primary qualities as much as to secondary. The fact that it is easier to describe a thought experiment in which someone has experience like ours except that chromatic color is missing than it is to imagine experience like ours without spatial features (if that is possible at all) does not affect the fact that an empirically contentful (as opposed, say, to a purely axiomatic) conception of space depends on visual or tactile or some other experience of a spatial field to give us a conception of what space might be empirically like, and that this is dependent on what it is like to perceive it in some particular way. P. F. Strawson (1959) argued that a purely auditory universe would not be enough to generate a conception of space, however the sounds were managed and organized. Whether he was right in thinking hearing alone could not generate a conception of space is not something we need now consider. What matters for present purposes is that, whether or not there could be a purely auditory and genuinely spatial world, we can certainly make *prima facie* sense of a mind with auditory experience, where the sounds are organized in a way which could not sustain a conception of space, and which lacks any other senses that might be sufficient to contribute a sense of space. The sounds it hears are simply those, say, of verbal discourse. If such a mind could be taught orally all the proofs of geometry and of relativity theory, it seems clear that its resultant grasp on the nature of empirical space would be no better than Mary's on color. This mind would have a purely scientific or formal conception, in a way that did not guarantee he would have any conception of what space was, or might be, like in

itself or qualitatively. I want to emphasize that any worries one might have about whether there could be such a mind as this are not to the point.² The *prima facie* intelligibility of the suggestion is enough to bring out the point that our conception of primary qualities as more than purely formal is not independent of what it is like to experience them. So, though the knowledge argument is most easily stated in terms of secondary qualities, which are seemingly easily relegated into the dustbin of the mind, the principle of the argument can be carried through for primary qualities that are fundamental for our conception of physical reality. Our conception of these is, at bottom, no more independent of what it is like to perceive them than is our conception of the secondary qualities. I say “at bottom” because primary qualities are not dependent on any particular form of experience, but this does not mean that we can have an interpreted or modeled conception of them without some form of qualitative content derived from experience. Furthermore, the doctrine of the ‘transparency’ of perception applies as much to primary as to secondary qualities, so what it is like to see square is only the obverse of what squareness is visually like.

Our own experience in fact bears this out. I suggested above that a mind with no spatial experience could not gain a better-than-abstract conception of space on the basis of learning geometry, relativity theory, and the like. But this is not so far from our situation. Insofar as our own grasp on four and above spatial dimensions, or on relativistic or bent space, is not purely mathematical, it depends on trying imaginatively to extend the two and three dimensions of which we have actual experience. This attempt is only very limitedly successful: we do not really achieve a grasp on what four and more dimensions, or other spatial exotica that go beyond experience, could actually be *like*. The narrative of *Flatland* makes plausible the thought that creatures that lived in two dimensions would have a similar difficulty to the one we have with more than three dimensions, in giving imaginative content to three dimensions.³ In none of these cases is the problem a lack of ‘theoretical knowledge’: it is a lack of the kind of experience that could give interpretation to that knowledge.

The correct way of looking at the rationale of the knowledge argument is to see it as granting content to the physicalist hypothesis

only for purposes of argument. “Even if we grant,” it says, “that physicalism could cope with the rest of reality, it still cannot cope with what it is like to experience things.” But once one recognizes the connection between what it is like to experience the world and what we can conceive the world we experience to be like, one can see that if physicalism cannot capture the former, it cannot capture the latter: one cannot have an adequate conception of the physical which does not include those qualitative components that are the ‘transparent’ projections of the qualitative nature of experience. Seen in this way, the knowledge argument begins to look as if it cannot fail to be right, for if there were not some special kind of content that is revealed only in experience, then we could not have an empirically significant conception of the physical in the first place. So, if you are tempted to think that physicalism might somehow be able to defuse the intuition that Mary learns something substantive and new, you need only direct your attention to the way that any nonformal conception of the physical is dependent upon the qualitative nature of reality as revealed uniquely in experience to see that this could not possibly be true. If, in general, the acquisition of experience did not teach something new, then a purely descriptive account of reality ought not to lack anything essential. In sum, the argument draws our attention to the fact that a physicalism that depends on a notion of the physical that is somehow independent of the qualitative nature of experience can present us only with a world that is so formal as to be empirically contentless.

How All Objections to the Knowledge Argument Miss the Point

A natural response to the argument so far might be as follows. It might be conceded that the KA, *if sound*, has a much stronger conclusion than has previously been thought. This merely emphasizes the need to show that it is not sound and increases the incentive to support one of the considerable set of objections that have been made to it.

This response is overoptimistic, however, from the physicalist’s perspective. All the responses to the KA of which I am aware assume that the physicalist’s conception of the material world is or could be

adequate for the nonmental realm and then explain how, starting from this basis, Mary's apparently new knowledge can be accommodated. I do not find these responses plausible, even in their own terms. But not merely does the KA challenge the physicalist's assumption that he has an adequate conception of the physical; it does so in a way that it is difficult to see that the physicalist, once the situation is drawn to his attention, can deny.

No one, I think, would wish to deny the following:

- (1) Our naïve, commonsensical or manifest image of the physical world essentially has qualitative features: that is, in addition to formal or mathematical features it has qualitative features which cannot be reduced to the formal ones.

From what has already been argued, it is clear that

- (2) These qualitative features derive, via the 'transparency' of perception, from the nature of qualia.

As qualia contribute an essential component in the commonsensical conception of the physical, it would seem that

- (3) The nature of qualia cannot be *analyzed* as some function of, or on, the operations of the physical, naively conceived.

This is relevant to the usual responses to the KA. Two of the most popular responses are the *abilities* response and the *phenomenal concept* response. Both these strategies take for granted the adequacy of a conception of the physical that does not essentially rely on the qualitative nature of experience to give it content, and then try to explain the latter—the qualitative nature of experience—in terms of this autonomous conception of the physical. In the case of the *abilities* account, the explanation of experience is in terms of behavioral abilities of physical organisms. In the case of the *phenomenal concept* strategy, it is in terms of a special form of conceptualization of certain physical states, nonmentally conceived. But if the qualia in experience are foundational for our notion of the physical, there is no au-

tonomous conception of the physical; so experiential states cannot be conceived in terms of some function on the physical as autonomously conceived.

This consequence might be taken as suggesting that the knowledge argument is set up in a way that is unfair to physicalism. It might seem to be unfair because it saddles the physicalist with having a purely descriptive or intellectual account of reality, and surely he is not denied the resource of sense experience in forming his conception: something must have gone wrong in our understanding of what physicalism or materialism requires.

Nothing, however, has gone awry. Of course, the physicalist is allowed to rely on perception to explain the acquisition of particular information about the physical world. But he is not allowed to draw essentially on the subjective dimension of experience—on what it is like to experience the world—in forming his conception of the physical nature of the world, for his conception is one committed to the availability of a purely objective account of the world. Insofar as the qualitative content of our conception of the world—that part which goes beyond what can be wholly captured descriptively—is a reflection of ‘what experience is like,’ it is a resource denied to the physicalist. This is the point at which traditional empiricism and physicalist realism as a metaphysical theory diverge. It is a starting point for empiricism that the qualitative components of experience are, or are among, the building blocks from which our conception of the physical world is constructed. Physical realists, on the other hand, simply ignore the role of perceptual experience, not simply in giving us information, but in giving our empirical concepts content. Whilst this can be thought of as an oversight, it is also essential to the orthodox physicalist project, for if the physicalist were to allow that *what it is like* to experience features of the world played an essential constitutive role in our conception of *what the world is like in itself*, he would have to abandon his fundamental project of assimilating the mental into, or reducing it to, the physical as autonomously conceived, for there is no such autonomous conception.

The way the argument undermines physicalism can be put even more comprehensively. A standard statement of physicalism is that it is the theory according to which phenomenal (and other mental)

states supervene with metaphysical necessity on physical states. Supervenience is an asymmetric dependence relation. It presupposes that the nature of the supervenience base does not essentially depend on that which supervenes on it: one can at least conceive of the base in the absence of the supervenient properties. But the KA shows that the physical cannot be conceived autonomously of elements dependent on the mental.

Now the physicalist might be tempted to argue that this is just a clash of intuitions: the proponent of the KA, as I interpret it, says there is no autonomous conception of the physical, and he, the physicalist, denies this. But I do not think this is true: the physicalist does not deny the role of the qualitative in any more-than-formal conception of the physical; he simply fails to notice the connection between this and the qualitative nature of experience and hence between it and qualia.

It is more or less explicit for the physicalist that

(i) we can have a grasp on the nature of the physical in scientific terms.

What I suspect most physicalists accept but which is not discussed in this context is

(ii) our concept of the physical is not purely mathematical and formal but involves a qualitative component.

Once one recognizes (ii), and one considers both or either of the KA (as I have re-expressed it) and the apparent transparency of at least some features of our experience, it is difficult to deny

(iii) we can have a clear grasp on a quality—be able to imagine what it is like—only if it is ultimately based on qualia: a quality is a ‘transparent’ projection or reflection of [some aspect of] a quale.

From this it follows

(iv) our conception of the physical is conceptually tied to or dependent on the nature of qualia.

From this it follows that

(v) qualia or qualia possessing states cannot be analyzed or explicated in terms of some function of or operation on the physical, as independently conceived.

2. THE NEUTRAL MONIST/TYPE-F MONIST RESPONSE

So, What's New?

I have talked as if there is at least a modicum of originality in my interpretation of the KA, but is it not essentially the same point as is made by Russell (1927), Maxwell (1978), Lockwood (1989), Galen Strawson (2006, 2007), and Stoljar (2006), who think that the scientific conception of matter is too abstract to accommodate consciousness and that this is what arguments such as the KA bring out? Certainly, I am moving in the same territory as Russell and these more recent philosophers, for we have in common the belief that a purely scientific physicalism fails to accommodate something which is essential to our overall conception of the world, namely those qualitative features of which we are consciously aware. This is in the vicinity of the theory labeled "neutral monism" by Russell and "type-F monism" by Chalmers (2003). It is valuable, however, to bear in mind the variety of tasks for which the qualities that standard physicalism cannot accommodate have been employed. My argument above has drawn attention to two.

(i) Without the special role of quality there could be no common-sense or manifest-image conception of the world, either in its primary or secondary quality aspects, and without this there could be no scientific image either. There could be no science without perception, no perception without sensible qualities, and no grasp on them without qualia, so even if there need have been no qualitative content in the subatomic world (as (iii) and (iv) below insist there must), quality must figure irreducibly in the world of experience.

(ii) Quite apart from the manifest image, our conception of the scientific world must include quality in its spatial properties. We might think we proceed as follows. We form a mental picture of the

web of causal powers that constitute the standard physicalist picture by imagining lines of influence and force similar to those by which we characterize a magnetic field. Then we ask ourselves whether this requires supplementing by qualities, as argued in (iii) and (iv) below. But the pure causal web is imagined in a (visual) qualitative space, even if it is imagined as an ontology of pure powers. In other words, quality comes in not just as a feature of the commonsense objects in space, but as an essential feature of the spatial medium itself, even in the scientific image. Space cannot be realistically conceived as purely mathematical, even in a world of pure energy and fields.⁴

These are the essential qualitative features that I have tried to indicate above, and they are concerned with our conception of the macro, not the micro, world. The post-Russellian tradition is mainly concerned with the role of quality at the micro level and how it can be deployed in the philosophy of mind.

(iii) Some philosophers, such as Armstrong (1968: 85–88) and Foster (2008: 71–72), have held that dispositional states must have categorical owners or bases. As science uncovers only structural and causal properties, these categoricals must be monadic and, in that sense, qualitylike.

(iv) The scientific account seems to construct the world from powers—forces, fields, energy—and there is a dispute about whether a world that consists purely of powers is incoherent. Those who claim that it is argue that powers to produce powers to produce powers . . . *ad infinitum* constitute a vicious regress. The point here is not that powers must be *owned* by something categorical—pure unowned powers or fields may be a possibility, according to this objection—but that they must, ultimately, result in some effect that is categorical. Again, the only clear candidate for this is something qualitative.⁵

(v) The neutral monist or type-F materialist project is to appeal to the qualitative nature of matter as a way of explaining the qualitative content of consciousness. In the world of consciousness, we are simply aware of the intrinsic qualitative nature of our brains, which science, as essentially the view from the outside of mere structural and relational properties, cannot reveal.

I am not convinced of the force of (iii)—maybe there can be unowned powers—but the necessity for qualities as specified in (i), (ii), and (iv) seems to me to be conclusive. But it is (v) which matters for

the philosophy of mind. The crucial issue in the philosophy of mind is whether the appeal to qualities deployed in any of (i) to (iv), which all purport to be, in some broad sense, features of the physical world, can be deployed in the articulation of a modified kind of materialism, which can be used to solve the mind-body problem.

I do not think that anyone would suggest that the possession of a qualitative nature by space could, on its own, at least, contribute to an analysis of the phenomenology of experience: so (ii) is not central to the issue. The qualitative nature of the manifest world is taken by some direct realists as helping to dissolve the mind-body problem: if the new quality Mary experiences when she leaves her room is a feature of the external world, then it is not an internal constituent of her mental state and so does not count against that state's being physical.⁶ This, as a strategy for reconciling experience and materialism, faces three problems. The first concerns the problematic nature of attributing secondary qualities to matter as intrinsic, mind-independent features. The second concerns the plausibility of direct realism, especially when charged with the task of being deployed to cope with *all* the kinds of qualities that we perceive and not just the obviously perceptual ones. Discussion of these two points would take me too far afield from the present discussion.⁷ The third difficulty is that direct realism of the kind being countenanced here is surely not a materialist theory. The relation between the perceiving subject and the objects and qualities he perceives would have to be a *sui generis* relation of awareness, and this is not part of a materialist ontology. Any attempt to replace this relation with something materialistically acceptable—say, some kind of purely causal relation—would leave one with a reductive account of experience: that is, if being aware of an external quality consists simply in the quality physically causing some physical process in the brain, that would render the presence of the quality phenomenologically irrelevant. It is not surprising, therefore, that it is the imputation of qualities to micro matter, as in (iii) and (iv), that has played a part in attempts to state type-F materialism.

Chalmers states this position as follows.

Russell pointed out that physics characterizes physical entities and properties by their relations to one another and to us. For example,

a quark is characterized by its relation to other physical entities, and a property such as mass is characterized by an associated dispositional role, such as the tendency to resist acceleration. At the same time, physics says nothing about the intrinsic nature of these entities and properties. Where we have relations and dispositions, we expect some underlying intrinsic properties that ground the dispositions, characterizing the entities that stand in these relations. But physics is silent about the intrinsic nature of a quark, or about the intrinsic properties that play the role associated with mass. So this is one metaphysical problem: what are the intrinsic properties of fundamental physical systems?

At the same time, there is another metaphysical problem: how can phenomenal properties be integrated with the physical world? Phenomenal properties seem to be intrinsic properties that are hard to fit in with the structural/dynamic character of physical theory; and arguably, they are the only intrinsic properties of which we have direct knowledge. Russell's insight was that we might solve both these problems at once. Perhaps the intrinsic properties of the physical world are themselves phenomenal properties. Or perhaps the intrinsic properties of the physical world are not phenomenal properties, but nevertheless constitute phenomenal properties: that is, they are protophenomenal properties. If so, then consciousness and physical reality are deeply intertwined. (Chalmers 2003: 130)

There are at least four problems with this type-F or neutral monist strategy, the fourth of which has not, as far as I can tell, received serious discussion. They are as follows. (i) How is one to move from the attribution of *quality* to matter, to endowing it with phenomenal *consciousness*? (ii) One must try to find a plausible account of what qualities or protoqualities (or phenomenal qualities or protophenomenal qualities) can be attributed to the elementary constituents of matter. (iii) How can one account for the unity of consciousness on the basis of the phenomenal/qualitative core of individual particles or events? The neglected one is (iv): More and more anti-reductionists seem to think that intellectual consciousness—states of conscious thinking—as well as sensory-type experiences, are irreducible. How can one apply the type-F strategy to these?

From Quality to Consciousness

At first sight, there would seem to be no reason why a qualitative core to atoms should provide any explanation of how consciousness emerges. Some defenders of neutral monism, for example Galen Strawson, try to solve the consciousness problem by adopting panpsychism. It is noteworthy, however, that earlier protagonists, such as Russell and Lockwood, thought that they could avoid compromising their physicalism in this way. For Russell, the qualities themselves are equivalent to the contents of *unconscious* mental states, and consciousness of them is given something close to a behavioral analysis: "A percept differs from another mental state, I should say, only in the nature of its causal relations to an external stimulus. 'Unconscious' mental states will be events compresent with other mental states, but not having the effects which constitute what is called the awareness of a mental state" (Russell 1927: 385). Consciousness is, in effect, quality plus appropriate effect on behavior. If one is not satisfied with this account of awareness, the problem of consciousness remains a major one.

Lockwood is not so behavioristic, but his original account of awareness seems to end up having, in its application of the concept of topic-neutrality to consciousness, more in common with Smart or Armstrong (1968) than one might expect from a neutral monist.

To the extent that we have a transparent grasp on the concepts that we bring to bear on our mental lives, those concepts may be seen as capturing certain intrinsic attributes of brain states. To the extent, however, that they are topic neutral, they represent no obstacle to an identity theory anyway. Moreover, *this goes for the concept of awareness itself*. For it seems to me that we cannot be said to have a transparent conception of awareness. . . . If that is right, then it follows that there can be nothing in our concept of awareness, such as it is, that could debar us from identifying awareness with some kind of physical process in the brain—albeit that it remains profoundly mysterious, in physical terms, what form such a process could possibly take. (Lockwood 1989: 169; emphasis original)

These are strange remarks. If our grasp on awareness is really topic-neutral, then it could be identical with any physical state that performs the right role. We may be ignorant of this, but there is no reason why its nature should be “profoundly mysterious.” Like many other philosophers, Lockwood seems to be confusing the fact that awareness is a simple, unanalyzable, and *transparent* relation with the idea that our concept is an empty one, waiting to be filled by some scientific theory. If the latter were true, then one could have a straightforwardly reductive account of mind, of a broadly causal or functionalist kind.

As Chalmers points out in the quotation above, neutral monism tries to kill two birds with one stone. A crucial gap is detected in our concept of matter, and this is remedied by deploying the concept of quality that derives from our ordinary experience to fill this gap. By importing these qualities, which might be thought of as qualia, neutral monism hopes to endow matter with the resources for generating experience. But if the qualitative element is genuinely neutral between mental and physical—‘merely qualitative’—then there is no explanation of how or why this should result in conscious states: we do not think of a red patch as being *per se* conscious. If one is to build the consciousness into the qualitative element that one is importing into the matter, then one has lost the neutrality and moved over to a panpsychism, which makes the core of matter at least minimally conscious. Russell and Lockwood seem to want to avoid doing that, but instead seem to end up with an account of consciousness which is somewhere between straightforwardly reductive and elusive. The source of the idea that the distinction between quality and consciousness can be blurred is, of course, Hume. The hope is that impressions are sufficiently phenomenal to be the building blocks for mind, without this phenomenality presupposing mentality as a principle in addition to their qualitative content. It is fairly clear, I think, that this cannot be done.

Imputing (Phenomenal) Qualities to the Fundamental Constituents of Matter

There are two ways of categorizing the fundamental constituents of matter. The most natural is to treat them as objects of certain kinds—

protons, electrons, quarks, gluons, and so on. These objects would have to have a qualitative core. The crudest version of this theory would think of these particles as consisting of little patches of color, or of sounds, itches, and so on, with these conceived of as not simply physical qualities but phenomenal or protophenomenal. The other option is to follow Russell and Maxwell in regarding the names of these so-called particles as names for groups of pure events. Maxwell thinks that the latter option is the only one that makes sense.

If C-fiber activity is thought of as consisting of threadlike pieces of matter . . . waving around and perhaps stroking each other, then any attempt to identify such activity with pain (as felt in all its excruciating immediacy) does become patently absurd. However, if we recognize that C-fiber activity is a complex causal network in which at least some of the events are pure events and that neurophysiology, physics, chemistry etc., provide us *only* with the *causal structure* of the network, the way is left open for the neuropsychologist to theorize that some of the events in the network *just are pains* (in all their qualitative, experiential, mentalistic richness). (Maxwell 1978: 386; emphasis original)

Maxwell does not explain why he thinks that the event ontology is so much more amenable to his theory, but one might imagine the reasons to be as follows. Pains are occurrent events, but C-fibres or complex brain states endure no doubt for a long time. If the pain is identical with the qualitative nature of the constituents of the fibers or cells, which overwhelmingly remain the constituents whether activated or not, why is the sensation not there for just as long? But it is not clear that an event ontology really gets round the problem. A brain process is, from the perspective of the subatomic, a massive and complex event. The pain must be a compound from the qualitative natures of the micro events that go up to compose it. But the ontology of particles cannot simply be ignored. Such things according to Russell, and of necessity if one has an exclusive ontology of events, must be compounds of events. The theory is not that it is a *mistake* to say that there are electrons, photons, and so on; rather, the theory is that such things are constructed from events. If we think of an

electron that endures from t_1 to t_n as constituted by events e_1 to e_n , then one might assume that, as the electron remains the same electron, its intrinsic nature does not change. Indeed, one might expect all electrons to have similar intrinsic natures. In this case there will have to be a consistency in the intrinsic nature of the events that compose it. The alternative appears to be that the qualitative nature of elementary particles—whether or not constructions from events—changes according to the nature of the large-scale causal net into which they are placed, and even though their causal contribution does not (electrons have always the same mass, the same charge, etc.). This would seem to be a very strange top-down phenomenon.

It is difficult to avoid the suspicion, from reading the passage quoted above from Maxwell, that he thinks of the pain event as a primitive, not something constructed from more primitive events.⁸ The same is true about Russell's (1927: 137) statement that in experience one is perceiving the inside of one's own head, as if the qualitative content of our sense data were what constituted the matter. There is, I suspect, an empiricist impulse to crunch together a phenomenalist and a physical realist conception of the world, without paying enough attention to the fundamental problems with this project. The introduction of the term 'protophenomenal' to characterize the qualitative nature of the elementary particles seems to me to name a problem, not to propose a solution. It amounts to no more than the suggestion that there must be something such that, if you get enough of it, you get a real experience, whilst hiding from the question of whether this involves moving from the non-experiential to the experiential. It is very difficult to form a conception of the consciousness of an earthworm—indeed to decide whether or not one can ascribe consciousness to it at all. What meaningful minimal consciousness-involving content is to be ascribed to a quark, or to one of the events that, as a group, constitute a quark?

The Problem of the Unity of Experience

On any type-F materialist strategy, the qualities which are supposed, in the end, to explain consciousness belong to the most elementary particles or events. When these qualities do constitute consciousness, they are bound together in unified sense-fields and in total

cross-media consciousness. How are we to explain this unity by reference to the phenomenal core of the individual elements? Notice that this is not the so-called grain problem, which is concerned with why a smooth and continuous consciousness should emerge from particles that are spatially distant from each other. The implication of the grain problem is that we should expect consciousness to be 'grainy'—a crude picture full of holes, gaps, and blanks. What I have called 'the unity problem' is the problem of explaining why there should be an overall picture at all, of whatever quality. This puzzle is strengthened if one considers what happens in other hunks of matter outside a limited area in brains. The lower brain, the kidney, and the table are all made up from matter which has a qualitative core, but no one seems inclined to attribute a unified subjectivity to them. If one followed Russell in treating consciousness as simply a matter of causal consequences, then there would be an answer to this problem, but such an approach to consciousness is no different from reductive functionalism.⁹ I quoted Lockwood above as seeming to adopt a reductive view, but in a later article he says that "it is difficult to see how *awareness itself* could be anything other than an emergent phenomenon" (Lockwood 1993: 280). Indeed, the bonding of the phenomena into a unity, though no doubt supervenient on functional organization, cannot be wholly explained by it, as liquidity is explained by atomic structure, and so must be emergent. But if awareness is emergent, what is achieved by attributing its objects to the matter of the brain? That awareness must be emergent is also attested by the following consideration. The individual elements have, at best, only the dimmest consciousness. Supposing them to be united into one consciousness, why should that not be equally dim? On what principle is the quality of consciousness accumulated?

Type-F Materialism and Intellectual States

The discussion of neutral monism is usually conducted with reference to the sensory *qualia* of consciousness, and most of the original protagonists of the theory were radical empiricists with reductionist accounts of thought. They tended to be imagists, associationists, or, later, behaviorists or functionalists about intellectual activity. More recent proponents of the theory tend not to share this reductionism.

Strawson, for example, rejects such reductionism, together with reductionism about sensory experience (1994: 4). The elementary entities, therefore, must possess not just protophenomenal qualities, but proto-intellectual content. This problem is easily overlooked because of the historical emphasis on the irreducibility of sensation rather than thought. Because perception gives us a conception of the physical world as being saturated with sensible qualities, it is natural for us to think of matter as essentially characterized by such qualities, and even, with some imaginative-cum-conceptual effort, by more primitive analogues of the same. The idea that minute matter manifests similar proto-intellectual features is harder to grasp. One might just about make a gesture in this direction by conceiving of the electron as possessing a protoconceiving of its own protophenomenal nature. This, however, will not be adequate to build up the distinctive character of thought unless all our thinking is built up logically from the concepts whose contents are restricted to the qualities they directly capture. This would be equivalent to a form of conceptual logical atomism, as found in linguistic phenomenalism. Such a program is both demonstratively impossible and presumably not what Strawson or the other contemporary Type Fs intend. One cannot build, without brute emergence, thoughts about Manchester United, the Trinity, or even our normal physical world from self-conceiving protophenomenal patches.

Explicit Panpsychism

Galen Strawson (2006, 2007), unlike Russell, Maxwell, and Lockwood, opts for the panpsychist solution. He calls his position “real physicalism,” but this label is misleading because ‘real’ does not qualify ‘physicalism’; rather the point is that the position is a physicalism that asserts the irreducible *reality of experience*. The use of ‘physicalism’ is also broad. Any concrete object that occupies space-time is physical, and this includes conscious animals such as ourselves. The bite in calling this ‘physicalism’ is a commitment to the idea that all the properties of such concrete objects, including the conscious states of those that are conscious, somehow flow from their nature as physical; that is, from the properties of the ultimate parts.

Strawson's argument for panpsychism can be reconstructed as follows (2006: 3–31).

- (1) Reductionism about experience is false.
- (2) Physicalism is true.
- (3) If reductionism about experience is false, and if physicalism is true, then, if “physical stuff is, in itself, in its fundamental nature, something wholly and utterly non-experiential,” there must be “brute emergence” of the experiential.

Therefore

- (4) If “physical stuff is . . . utterly non-experiential,” then there must be brute emergence of the experiential.
- (5) Brute emergence is an incoherent idea.

Therefore

- (6) Physical stuff is not in its fundamental nature utterly non-experiential.

The argument is valid. Strawson regards (1) as intuitively obvious, but it is also *ex hypothesi* at the current state of the argument in this paper. The strategic situation is that, if (5) is true, either matter is essentially experiential or one must abandon physicalism and accept that consciousness is an essentially different nature or substance from matter, as the dualist claims. Why is Strawson so convinced of (5)? There are, I think, two reasons. First, he argues that what are usually cited as cases of emergence—for example, liquidity arising from atoms that are not themselves liquid—are not cases of brute emergence, because the nature and the behavior of the atoms rationalize and entail the liquid product. Once you understand how the atoms behave, you are in a position to see that the macro phenomena could not fail to be liquid. Second, he thinks that the belief that there could be such a thing as brute emergence derives from an exaggerated conception of what it is for causal relations to be contingent, which itself derives from a misunderstanding of Hume: once one is purged of this error, then one will see that brute emergence is a nonsense.

I cannot here discuss causation in general or the interpretation of Hume, but it is generally accepted that the emergence of consciousness involves an ‘explanatory gap’ not present in cases such as liquidity. It therefore follows for the physicalist that either the seeds of experience are in matter in the panpsychic sense, or experience emerges in so unique a way as might be thought to constitute a form of dualism.

Suppose we agree that a physicalist who accepts the irreducible reality of experience is obliged to be a panpsychist. Why should we not regard this as a *reductio* of physicalism rather than an argument for panpsychism? Why, in other words, does Strawson think that it is so plausible to claim that physicalism can absorb the experiential? The answer is that he thinks that, given the topic-neutral nature of our scientific conception of matter, physicalism that accommodates irreducible experience is common sense. He approvingly quotes Eddington, who is asking whether there is anything in our knowledge of matter that prevents attributing mental properties to it.

In regard to my one piece of insight into the background, no problem of reconciliation arises; I have no other knowledge of the background with which to reconcile it. . . . *There is nothing to prevent the assemblage of atoms constituting a brain from being a thinking [conscious, experiencing] object in virtue of the nature which physics leaves undetermined and undeterminable.* If we must embed our schedule of indicator readings in some background, at least let us accept the only hint we have received as to the significance of the background—namely that it has a nature capable of manifesting itself as mental activity. (Eddington 1928: 259–60; quoted in Strawson 2006: 11; emphasis original)

Strawson believes that the experiential and the scientific conception of the physical slot easily together and that treating them dualistically pointlessly offends against Ockham’s razor. This of itself, of course, does nothing to answer the four problems I raised above for type-F monism, so how does Strawsonian panpsychism fare on these?

The first problem—how one moves from qualitative content to consciousness—does not arise for a theory that is explicitly panpsy-

chist. On the second problem—the mental life of quarks and strings—I see no helpful guidance in Strawson.

This problem is to say what kind of mental life a quark or a proton is supposed to possess. There are two forms of panpsychism, and Strawson's physicalism commits him to the less plausible. The more plausible version is holistic, in that it sees the whole of the material universe as somehow pervaded or infused by mind or intelligence. This 'world spirit' is a property of the whole and is not constructed from the mental features of the parts. The other version is atomistic—Strawson calls it 'smallism'—and seeks to attribute to each atom an appropriately minute form of consciousness and to build more sophisticated consciousnesses out of this material. The mentality in the former case, though no doubt mysterious, is at least modeled on mind, spirit, and intelligence as we know it. In the latter case, it is utterly obscure what the atomic materials could consist in. When setting out this problem, I said that the consciousness of an earthworm—a massively complex organism by subatomic standards—is hard enough to conceive and asked what it might mean to attribute a suitably diminished version of consciousness to an electron. I suggested that it might be impossible to imagine or give theoretical content to such an idea. It is important that it is not just a matter of imagination, in the way that it might be impossible to imagine what it is like to be a dog, even though one is quite confident that there is something that it is like. It is plausible to maintain that being conscious involves a certain complexity of structure: one is taking something in a certain way and responding to it. This is true of a dog and just possibly of an earthworm. 'Responding' need not mean external behavior: it could be any mental affective or cognitive response. But the occurrence of a single, minimal qualitative content, in association with an undifferentiated external causal response (for example, in the case of an electron, unalterably exercising the influence of the mass of $1/1860$ of a proton and of a negative charge), cannot constitute any inner consciousness. It seems to me reasonable to think that the existence of any kind of subject presupposes a certain movement of mind and hence the active grasp on more than one content. Strawson says that the experiential is always active, not passive, but the only activities that he can ascribe to the electrons, strings, and so on

are the kinds of external, unvarying properties I have cited for the electron: this is not mental activity.

Strawson has more to say about the unity of nonsimple consciousnesses, which he calls “the composition problem.”¹⁰ There are in fact three components to this problem. First, there is the issue of why and how the inner core of separate simples could or should merge into one consciousness. Second is the issue of why and how the very dim and different contents of simples come to make up the kind of conscious experiences we have, given that they can come together at all. Third is the issue of how separate *subjects* can make one subject, especially whilst escaping detection to introspection.

It is not clear to me that Strawson distinguishes the first two problems. On the first problem, he cites William James as having gone from believing that composition was impossible to believing in a “not-rigidly-particulate, field-quanta-friendly form of Composition” (Strawson 2006: 248). The rationale is as follows. Provided that one has a field, rather than a particulate, conception of the simples, the thought that they overlap and ‘flow into’ each other to form a new unity does not seem so unnatural. Furthermore, this helps with the second problem, because this intermixing might explain how they can produce contents that are significantly different from those of the elements that mix—their fusion is more like something chemical than like mere physical combination. These thoughts seem to be backed up by two general principles. One is the optimistic belief that “we know it is actual so it must be possible,” and the other is that “unintelligible experiential-from-experiential emergence is not nearly as bad as unintelligible experiential-from-non-experiential emergence” (Strawson 2006: 250).

We have now moved a long way from the original position, where the role of the qualitative/mental was to provide the monadic intrinsic properties of matter, to which the causal properties discoverable by physical science could belong. The mental atoms have now developed a chemistry of their own which does not seem to follow from the physical laws which were originally conceived to be their only powers. One has disposed of the obligation to make any sense of how or why the mental developments come from their elements. We are much nearer to a holistic idealism than we were at the start of the

project. Perhaps one is near to thinking of the whole process as having some mind-serving teleological focus.

The third problem of composition was how many selves can form one. This itself has two subproblems. One of these, raised by Goff (2006), is how, given the transparency of consciousness, we could fail to notice that we were constituted “like the eye of a fly.” Strawson’s reply to this specific point seems to me to be adequate. He says that transparency of consciousness does not guarantee awareness of all its features. I would put it by saying that we are aware of qualitative content but not metaphysical structure. As an analogy, it could be pointed out that simple introspection does not reveal the correct philosophical ontology of perception, only its qualitative phenomenology. The more serious problem, I think, concerns how one subject can be aware of the logically private contents of another mind. Either what I am aware of is, in some fused way, identical with the contents of all the lower-level subjects, or it is a causal product of these. The former option infringes logical privacy; the latter is a case of brute emergence. Strawson prefers conscious-to-conscious brute emergence over unconscious-to-conscious, but this seems to me a pretty desperate position.

On the fourth general problem—the emergence of thought—the original problem remains for any version of panpsychism that seeks to have a tight and systematic account of how thought can develop from minimal qualitative consciousness. Whether the allowance of Jamesian flowings and fusings really makes this any better is hard to estimate.

The Appeal to Ignorance

Daniel Stoljar (2006) gives what might seem to be an alternative to Strawson’s panpsychism. Stoljar’s starts from the Russellian position that there is something about the nature of matter that current science cannot tell us and of which we are ignorant. He does not claim that this gap is filled by the qualities revealed in experience, but he does believe that it is filled by something which, were we to come to know its nature, we could see how it gives rise to experience. Stoljar, therefore, circumvents Strawson’s direct move to panpsychism,

whilst accepting his rejection of brute emergence in favor of the belief that the underlying nature of matter must be able to provide an explanation of the development of consciousness. We are just ignorant of what this underlying feature is. He believes that this appeal to ignorance, which has other parallels in the history of science, is more plausible than reifying our ignorance into dualism.

Stoljar recognizes that the main opposition to the suggestion that we are simply ignorant of what it is about matter that enables it to produce consciousness comes from the conviction that, in principle, nothing that was any sort of physical feature could constitute an explanation of mentality. And he recognizes that the principal reason behind this conviction is that, to be physical, a feature must be objective, that is, equally available to anyone, and that no such feature could explain subjectivity—which consists in features available in a special way only to the subject who has them (Stoljar 2006: 153–62). The crucial part of his positive argument, therefore, consists in his attempt to show that this divide can, in principle, be crossed. He has a two-pronged argument for this conclusion.

First, Stoljar argues by counterexamples (2006: 157–62). He argues that

(i) John is in pain

is a subjective statement, and that, therefore,

(ii) John is not in pain

is also subjective because it “contains the same constituents” as (i).

Moreover,

(iii) John is a number

is an objective statement, but it is true that

(iv) if John is a number then John is not in pain

is a true entailment. So statements with objective subject matter can entail statements with subjective subject matter.

In case one finds something linguistically odd about (iv), he runs a parallel argument through for

(v) if John does not exist then he is not in pain,

to which the same objection cannot be raised.

The initial reaction to this line of argument is to be suspicious of the way it deploys negatives. The fact that certain categories necessarily exclude each other seems not to throw light on whether they might positively entail each other, and if so, how. This suspicion is correct, for this negative strategy, if sound, would prove far too much. It is true that

(vi) if seven is a number then it is not spatially extended

and

(vii) if John does not exist he is not spatially extended.

It would be bizarre, however, to take these propositions as giving any kind of support to the hypotheses that there might be a feature of abstract objects or a feature of nonexistents of which we are currently ignorant which could explain how some or all of them might actually possess, or have possessed, spatial or other physical properties.

In fact Stoljar's principle of argument seems to entail the following absurdity.

(viii) If an object's belonging to category X (possessing X-type properties) logically excludes its possessing Y-type properties, then it follows that another object belonging to category X might possess properties that entail that it does possess Y-type properties.

But it surely cannot be right that the fact that one object in a certain category, X, is conceptually excluded from belonging to another category Y entails that there must be something else in X which is not so excluded.

What this shows is that entailments of exclusion do not throw light on the possibility of positive entailments.

Second, Stoljar presents a diagnosis of why we mistakenly think that no objective fact could ever seem fully to explain experience (158–62). He thinks that we confuse two things. ('N' represents the physical feature of which we are currently ignorant, and 'E' the experiential state.) They are

(1) Even if you were to know N, you would still not thereby know E.

And

(2) Even if you were to know N, there would still appear to you to be an element of contingency in the relation between N and E.

The mistake is thinking that (2) follows from (1), whereas (1) is true and (2) is false. The issue turns on whether "if N then E" is what Stoljar calls *synthesizable*, which it is if and only if understanding N involves understanding E. It is because that conditional is not synthesizable that (1) is true. It does not follow, however, that the relation would appear to be contingent, as claimed in (2).

I think that this diagnosis is false. Synthesizability seems to be more or less equivalent to definability of E in terms of N, because it requires that understanding of N is itself sufficient for understanding E. This contrasts with the case in which you independently understand both and can see how N is sufficient for E. Taking 'A' to stand for the atomic structure that makes something a liquid,

if A then x is a liquid

is not synthesizable, for someone might, theoretically, have the vocabulary of atomic or molecular science without having the concept of a liquid. One can nevertheless see how A is adequate for constituting something as a liquid. But this is not the problem with experience. Before the chemistry of liquidity was uncovered, sophisticated people would not have thought that liquidity was the kind of phenomenon for which no possible physical and mechanical constitution could be found. Standard scientific reductions are not synthesizable, but that did not create resistance to the prospect.

In short, it seems to me that Stoljar fails to undermine the intuition that no kind or kinds of objective facts could be seen as sustaining experiential states in a way analogous to that in which facts about atoms can be seen as sustaining facts about liquidity. But furthermore, the point about intellectual states I raised against Strawson applies against Stoljar. Even if some unknown feature could rationalize in a bottom-up way the emergence of sensory experiences, can we really believe that it could do the same for our intellectual states, however subtle or abstract? Stoljar does discuss intellectual character, but only in the context of Descartes's antiphysicalist argument that a machine could not think *because it was inconceivable that its behavior should show the appropriate subtlety* (124–25). Stoljar cites this as a historical parallel to his own argument, because later science revealed that matter does possess the relevant property. This suggests that Stoljar, like Russell and, earlier, Lockwood, but unlike Strawson, is satisfied with a reductionist (in this case, a computer functionalist) account of thought. This essentially ignores the phenomenology of thought.

Stoljar, therefore, in my opinion, fails to vindicate type-F materialism by an appeal to our ignorance.

■ One might summarize the argument of the second part of this paper as follows.

The knowledge argument shows

(1) there can be no adequate account of the conscious mind in standard physicalist terms.

It follows from this that, unless one accepts a uniquely brute form of emergence,

(2) if mind is to be explained from a materialist perspective at all, there must be some feature of matter in addition to those contained in a standard physicalist account which, unlike the standard ones, does provide an explanation of the generation of the conscious mind.

(3) This further feature must itself be either mental or protomental (conscious or protoconscious)

because

(4) nothing that was purely and simply physical—nothing that was essentially accessible from a third-person perspective—could conceivably explain the generation of the subjective and hence the mental.

This last is what Stoljar denies, but I have tried to refute his arguments.

(5) The attribution directly to elementary matter of full-fledged mental properties, the same kind of properties as figure in our experience—colors, sounds, itches, and so on—is totally bizarre. This is even more especially true in the case of the contents of intellectual conscious states.

Maxwell, we have seen, tried to get around this problem by locating these qualities (the sensible, not the intellectual, which he does not mention) at a relatively macro level, but Lockwood shows that this makes his theory a standard version of emergentism.

Therefore,

(6) the materialist needs proto-mental, or proto-conscious, states.

But

(7) no clear sense has been given to the notion of such proto-mental or proto-conscious states that differentiates them from *whatever it is* that adequately explains the generation of mind. It does not help in understanding what sort of thing might provide such an explanation, what such a thing might be like, or that there could be such a thing.

Furthermore,

(8) there is no remotely plausible account of how proto elements might combine to produce full or normal conscious states. One would be forced back to emergence. This is especially true of intellectual states.

Overall, the theory-based physicalism of the physicist cannot capture the qualitative nature of the world and so is condemned, as a total worldview, to be incoherently abstract. This cannot be remedied by any version or development of neutral monism, which tries to load what is missing in the physicist's world into a richer conception of matter. If there is a physical world independent of our experience, it cannot provide an explanation of why that experience should exist. There cannot be a materialist, or a materialistically based, monism that is adequate to the phenomena.

NOTES

1. The literature on the knowledge argument is vast, but classic pieces can be found in Ludlow et al. 2004.

2. In case one is worried by the apparent impossibility of such a case, one might consider the following. Imagine someone who had developed with normal spatial experience, but then suffered brain damage that destroyed all memory of the spatial features of his experience whilst not harming his general and mathematical intelligence. There was then an attempt to teach him scientifically the properties of space. This would lead to the same situation.

3. See Abbott 1884/2006.

4. For a demonstration of this, see Foster 1982: 176–88.

5. The dispute about the powers conception of reality has a growing literature. Examples are Robinson 1982: 108–23; Robinson 2009; Foster 1982: 67–72; Blackburn 1990; and Molnar 2003.

6. This general line has been strongly defended in discussion by my colleague Hanoch Ben-Yami. I am very grateful for his contribution to the development of my ideas in this paper.

7. They are discussed in Robinson 1994.

8. Maxwell remarks that perhaps it is “this ‘middle sized’ realm that provides the relevant context for investigation of mind-brain identities” (1978: 399). Lockwood takes this as suggesting that the neutral monism

applies only at a relatively macroscopic level, and shows that such a theory is no different from a standard sort of emergence (Lockwood 1993: 280–81).

9. Some empiricists—for example, A. J. Ayer—seem to think that it is possible to be non-reductionist about qualia, but reductionist about our cognition of them. For an argument that this is impossible, see Robinson 1982: 105–7.

10. For discussion of this issue, see Strawson 2006. The discussion by Goff in that volume is very clear and helpful. Strawson’s reply is at 248–52.

REFERENCES

- Abbott, E. A. 1884. *Flatland: A Romance of Many Dimensions*. London: Seeley and Co. Repr., Oxford: Oxford University Press, 2006.
- Armstrong, D. M. 1968. *A Materialist Theory of the Mind*. London: Routledge & Kegan Paul.
- Blackburn, S. W. 1990. “Filling In Space.” *Analysis* 50:62–65.
- Chalmers, D. 2003. “Consciousness and Its Place in Nature.” In *The Blackwell Guide to the Philosophy of Mind*, edited by Stephen Stich and Ted Wardfield, 102–42. Oxford: Blackwell.
- Eddington, A. 1928. *The Nature of the Physical World*. New York: Macmillan.
- Foster, John. 1982. *The Case for Idealism*. London: Routledge & Kegan Paul.
- . 2008. *A World for Us*. Oxford: Clarendon.
- Goff, P. 2006. “Experiences Don’t Sum.” In G. Strawson et al., *Consciousness and Its Place in Nature*, edited by A. Freeman, 53–61. Thorverton: Imprint Academic.
- Hume, D. 1978. *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge, revised P. H. Nidditch. Oxford: Clarendon.
- Jackson, F. 1982. “Epiphenomenal Qualia.” *Philosophical Quarterly* 32:127–36. Reprinted in Ludlow et al., *There’s Something about Mary*, 39–50.
- Lockwood, M. J. 1989. *Mind, Brain and the Quantum*. Oxford: Blackwell.
- . 1993. “The Grain Problem.” In *Objections to Physicalism*, edited by H. Robinson, 271–91. Oxford: Clarendon.
- Ludlow, P., Y. Nagasawa, and D. Stoljar. 2004. *There’s Something about Mary: Essays on Phenomenal Consciousness and Frank Jackson’s Knowledge Argument*. Cambridge, MA: MIT Press.
- Maxwell, G. 1978. “Rigid Designators and Mind-Brain Identity.” In *Perception and Cognition Issues in the Foundations of Psychology*, edited by C. Wade Savage, 365–403. Minnesota Studies in the Philosophy of Science 9. Minneapolis: University of Minnesota Press.
- Molnar, G. 2003. *Powers: A Study in Metaphysics*. Oxford: Oxford University Press.

- Robinson, H. 1982. *Matter and Sense*. Cambridge: Cambridge University Press.
- . 1994. *Perception*. London: Routledge.
- . 2009. "Idealism." In *The Oxford Handbook to the Philosophy of Mind*, edited by Brian McLaughlin, Angsar Beckermann, and Sven Walter, 189–205. Oxford: Oxford University Press.
- Russell, B. 1927. *The Analysis of Matter*. New York. W. W. Norton.
- Stoljar, D. 2006. *Ignorance and Imagination: On the Epistemic Origin of the Problem of Consciousness*. Oxford: Oxford University Press.
- Strawson, G. 1994. *Mental Reality*. Cambridge, MA: MIT Press.
- . 2006. "Realistic Monism: Why Physicalism Entails Panpsychism" [keynote paper] and reply to commentaries. In G. Strawson et al., *Consciousness and Its Place in Nature*, edited by A. Freeman, 3–31. Thorverton: Imprint Academic.
- . 2007. *Real Materialism, and Other Essays*. Oxford: Oxford University Press.
- Strawson, P. F. 1959. *Individuals*. London: Methuen.