

## NATURALISM AND PHYSICALISM

Encyclopedia entry in Robert Barnard and Neil Manson, eds., *A Companion to Metaphysics*, Continuum Publishing, 2012.

### Two influential 'ism's

Over the last several decades a great deal of work in metaphysics has been motivated by an allegiance to views most commonly known by the labels 'naturalism' and 'physicalism' (sometimes 'materialism'). The influence of these commitments is most easily seen in work in the philosophy of mind, but both doctrines aim to have more general import, constraining one's overall view of the world.

Exactly what these labels are meant to denote is not always clear, and this problem is exacerbated by an occasional tendency to use them more as a badge of respectability than as a useful indication of philosophical commitments. The two terms are certainly not used interchangeably. One obvious difference is that 'naturalism' is often deliberately used to name something other than a metaphysical thesis, signifying instead a commitment to a method of inquiry, perhaps accompanied by a theoretical account of that method's status. Call this 'epistemic naturalism.' By contrast, 'physicalism' is normally used only to signify a metaphysical thesis. The two terms differ, further, in that physicalism seems to be a thesis logically stronger than (metaphysical) naturalism. The label 'naturalist dualist' is not only intelligible but actually claimed by some philosophers (e.g., Chalmers, 1996, pp. 124-129; see also the 'panpsychist naturalism' of Sellars, 1927). By contrast, it is harder to see how one could make sense of a declared anti-naturalist physicalist, though the combination sounds intelligible if the naturalism rejected is a purely epistemic thesis independent of the metaphysical physicalist claim.<sup>1</sup>

The present entry is divided into three main sections. In the first, after a very brief discussion of epistemic construals of naturalism, I survey some of the ways in which metaphysical naturalist might be understood. I also suggest a general interpretation of 'naturalist' talk, the 'nothing special' interpretation, according to which adherents of metaphysical naturalism are motivated by the core idea that humanity and our interests are not special in the overall order of things. This formulation does not yield a determinate metaphysical thesis but does provide a schema into which such theses may be fit—including physicalism as a version of naturalism.

The second section is devoted to physicalism and its formulation. The simple thesis that everything is physical will not be adequate, as there are recognizably physicalist positions that allow for entities which are strictly speaking not physical but which bear some appropriate relationship to the physical. More precisely, such positions allow for entities such that their existence and character is determined by the physical in a way that makes it appropriate to say that those entities are 'nothing over and above' the physical entities involved in that determination. A thorough understanding of physicalism will require attention to this relation, but there are two other parameters that need

---

<sup>1</sup> If epistemic naturalism forbids a priori knowledge, then someone like Frank Jackson (Jackson, 1998), who embraces both physicalism and a priori knowledge, would count as a kind of anti-naturalist physicalist. I would count myself in this camp as well (Witmer, 2006).

clarification as well. Exactly what does it mean to describe an entity as 'physical' in the first place? And just how extensive is the reach of the physicalist claim? That is, just what does the 'everything' range over in the claim that everything is appropriately related to the physical? (Does it include, for instance, abstract objects?)

The third and final section addresses the ways in which one might justify either of the two doctrines. Challenges to them are well-known; with regard to physicalism, the most prominent prima facie problem is to accommodate mentality. In this article I only address the question of how the claims may be supported in the first place. Given the widespread allegiance to these doctrines and the often casual way in which they are invoked in contemporary philosophy, it is easy to suspect that there is more dogma than argument behind their endorsement. Nonetheless, we can identify a number of important lines of support, all of which turn in some way on the success of the natural sciences, especially physics.

### **Interpreting 'naturalism'**

In the introduction to his book Philosophical Naturalism, David Papineau writes:

What is philosophical 'naturalism'? The term is a familiar one nowadays, but there is little consensus on its meaning. For some philosophers, the defining characteristic of naturalism is the affirmation of a continuity between philosophy and empirical science. For others the rejection of dualism is the crucial requirement. Yet others view an externalist approach to epistemology as the essence of naturalism.

I shall not engage directly with this issue. It is essentially a terminological matter. The important question is which philosophical positions are right, not what to call them. I suspect that the main reason for the terminological unclarity is that nearly everybody nowadays wants to be a 'naturalist', but the aspirants to the term nevertheless disagree widely on substantial questions of philosophical doctrine. The moral is that we should address the substantial philosophical issues first, and worry about the terminology afterwards. (Papineau, 1993, p. 1)

While Papineau is right to stress the substantive question over the terminological one, it may nonetheless be a good idea to have a plausible account of what inclines us to think of one position as 'naturalist' and another not, since that may help us uncover just what makes the positions categorized as naturalist seem attractive to so many. With such an account on hand, we may be able to see good reasons for or against substantial positions that were not previously open to view.

As noted above, 'naturalism' is sometimes used to denote either an epistemic thesis or a commitment to a method of inquiry; this is illustrated in the quick gloss offered by Papineau. After a very brief discussion of such I turn to our main focus: metaphysical naturalism.

#### Epistemic naturalism

Epistemic naturalism is associated with two familiar slogans:

- (1) There is no first philosophy.
- (2) Philosophy is continuous with the sciences.

The first slogan may be taken as the claim that philosophers should not try to arrive a priori at conclusions that are then used to evaluate, limit, or otherwise regulate non-philosophical inquiries. While the first slogan denies, the second affirms: there is a legitimate role for philosophy in our investigation of the world, but there is nothing distinctive or special about its methods.

Exactly what would it take for philosophy to be discontinuous with the sciences? Just what does (1) prohibit? It might be taken as a ban on all a priori theorizing, and this is certainly how many self-proclaimed naturalists see their commitments, though this is not true of all of them (Goldman 1999). Such naturalists often name W. V. O. Quine as an important influence, and Quine of course famously denies the existence of a priori knowledge. Perhaps the most commonly cited work is Quine's 'Epistemology Naturalized' (Quine, 1969), the main focus of which is not a denial of a priori knowledge but his recommendation that we make free use of empirical science in theorizing about the relations between theory and evidence, and this recommendation is in turn driven by his rejection of a certain project, namely, that of using philosophical argument to legitimate the pursuit of scientific inquiry and its results. One way philosophy might be distinctive, then, is if it were capable of properly pursuing that project. The epistemic naturalist rejects this idea: philosophy has no special resources that would enable it to take this lofty perspective on other areas of inquiry. As Penelope Maddy puts it, the 'fundamental naturalistic impulse' is 'a resolute skepticism in the face of any "higher level" of inquiry that purports to stand above the level of ordinary science' (Maddy, 2001, p. 39). She cites Quine approvingly as ruling out any 'supra-scientific tribunal' that would be in a position to 'justify or criticize science on extra-scientific grounds' (Maddy, 2001, p. 43).

Why should 'naturalism' be used as a term covering both this epistemic commitment and a metaphysical thesis that rules out such things as ghosts and gods—things that are 'supernatural'? If one is convinced that the only way philosophy could play the special legitimating role would be if philosophers had access—putting it bluntly—to magic, then it is easy enough to see the connection. (Critics of a priori knowledge often speak of alleged 'occult faculties' that would underwrite such knowledge.) Of course, one might think that philosophy can play such a role without pretending to be magical (whatever, exactly, being magical is supposed to come to), and in that case one might find the common nomenclature inappropriate.

### Metaphysical naturalism

Just how should the metaphysical thesis be understood? 'Naturalism' (hereafter to designate the metaphysical claim) is plainly meant to rule out the existence of certain

unwelcome entities—those characterized as the ‘supernatural’ or the ‘non-natural’.<sup>2</sup> We might take (3) as a suitable starting point:

(3) Every entity is a natural entity.

This does not help much until we have some grip on what it is for an entity to be ‘natural’, of course.<sup>3</sup>

Many remarks about metaphysical naturalism lean on examples of the unwelcome entities in order to fix the thesis at issue. For instance, John Dupré offers us the following:

By antismetaphysicalism I mean something like the denial that there are entities that lie outside of the normal course of nature. It is easier to point to some of the things that are agreed to lie outside the normal course of nature than it is to characterize the normal course of nature. Central cases of such outliers are immaterial minds or souls, vital fluids, and deities. (Dupré, 2004, p. 36)

Mario De Caro lists several examples of things contrary to naturalism: ‘spiritual entities, Intelligent Designers, immaterial and immortal minds, entelechies and prime movers unmoved’ (De Caro, 2010, p. 367). As De Caro points out, however, these examples don’t help much once we move ‘beyond the simple cases toward more problematical ones, such as values, abstract entities, modal concepts, or conscious phenomena’ (367). Are those supernatural or non-natural? How are we to tell? Certainly it won’t help to ask if they involve ‘magic’.

A remark by Barry Stroud aims at something more general:

By ‘supernaturalism’ I mean the invocation of an agent or force that somehow stands outside the familiar natural world and whose doings cannot be understood as part of it. (Stroud, 2004, p. 23)

Dupré invokes the ‘normal course of nature’; Stroud here speaks of something being apart from the ‘familiar natural world’. But appeals to familiarity or regularity don’t seem quite appropriate. There seems nothing odd about the idea that nature may contain very unfamiliar elements; indeed, avowed naturalists are often insistent on the potential

---

<sup>2</sup> There is a nice question as to how ‘supernatural’ and ‘non-natural’ are to be distinguished from each other. The latter seems designed to be less offensive than the former, with perhaps the ‘super’ indicating precisely the same unwanted aspect as does the ‘supra’ in Maddy’s talk of a ‘supra-scientific tribunal’—that is, some kind of authority or power over the merely natural.

<sup>3</sup> There is a use of ‘natural property’ in contemporary metaphysics that must be distinguished from the use of ‘natural entity’ in the current context. As used by David Lewis (Lewis, 1983) and many others influenced by him, the natural properties are those such that their instances genuinely have something in common, so that positing a universal is tempting, whereas the non-natural properties are merely classes of possible individuals, where they may have nothing genuinely in common. Plausibly, if we say that some predicates correspond to something in reality other than the individuals that satisfy them while others don’t, the former pick out the ‘natural’ properties and the latter don’t. To see that this notion of ‘natural property’ is not what is currently at issue, consider the property of being a ghost. If there were ghosts, then that property would surely count as natural in the Lewisian sense, but presumably it would still be objectionably supernatural in the sense at issue in metaphysical naturalism.

strangeness of the natural world.<sup>4</sup> Nor does it seem out of keeping with our intuitive grasp of the natural to say that some natural events may be very infrequent (think of the Big Bang) or irregular (think of quantum indeterminacy).

One well-known use of 'naturalism' as a metaphysical thesis is due to David Armstrong, who offers an account that doesn't lean on examples of this sort. He defines naturalism as the claim that '[t]he world is nothing but a single spatio-temporal system' (Armstrong, 1978, p. 126). Let us then consider (4):

(4) Every entity is located in a single spatiotemporal system.

In the same work, Armstrong explicitly distinguishes 'materialism' from naturalism, describing materialism as the thesis that 'the world is completely described in terms of (completed) physics' (126). If we presume that a completed physics will only posit things that are located in a single spacetime but that some possible spatiotemporal entities would not be included in such a physics, we have the result that materialism is logically stronger than naturalism, given Armstrong's definitions.

But (4) is not just weaker than materialism; it is surely weaker than what advocates of naturalism have in mind. Being located in space and time does not, on its own, do very much to constrain the kinds of items that might be thus located (see also Kim, 2003). It is far from clear, for instance, that (4) rules out the existence of ghosts as traditionally conceived. Such things have very odd and surprising properties, to be sure, but the man on the street would have no trouble thinking of ghosts as existing at particular places and at particular times. Nonetheless, it is safe to say that avowed naturalists would reject such entities as belonging to the category of the 'supernatural'.<sup>5</sup>

One can find throughout the writings on naturalism an emphasis on spatiotemporal location—and, often, an emphasis on a 'space-time-causal system'.<sup>6</sup> For a contemporary example, consider how George Gasser introduces a recent anthology of papers on naturalism. He first describes the epistemic naturalist thesis as the claim that 'all knowledge we can acquire is obtainable only or foremost through the application of the scientific method' (Gasser, 2007, p. 4) but follows this with:

---

<sup>4</sup> For example, when James Ladyman and Don Ross explain why it is 'anti-naturalist' to give weight to intuitions, they claim that doing so 'requires ignoring the fact that science, especially physics, has shown us that the universe is very strange to our inherited conception of what it is like' (Ladyman and Ross, 2007, p. 10).

<sup>5</sup> Only one page after his account of 'naturalism' Armstrong says that to accept naturalism 'is to reject such entities as Cartesian minds, private visual and tactual spaces, angelic beings and God' (Armstrong, 1978, p. 127). Armstrong presumably can appeal to his official doctrine (4) to rule out each of these. Cartesian minds, angels, and God are presumably not located in both space and time, while private visual and tactual spaces may be located in some way but not in the same spatiotemporal system as ordinary objects. Armstrong's emphasis on a single spatiotemporal system thus does play a role. Unhappily, however, this same emphasis renders the thesis too strong. Suppose the world were to consist of two distinct spatiotemporal systems, each containing nothing but what is describable by current physics. I doubt any naturalist would decry such a world as contrary to his doctrine simply because of the disconnect between two otherwise perfectly 'natural' systems.

<sup>6</sup> For example, consider Roy Wood Sellars' remarks in 1927: '[N]aturalism takes nature in a definite way as identical with reality, as self-sufficient and as the whole of reality. And by nature is meant the space-time-causal system which is studied by science and in which our lives are passed' (Sellars, 1927, p. 217).

This epistemological thesis often comes hand in hand with an ontological thesis claiming that all that exists is what (in principle) can be studied by science. Science studies the spatio-temporal world. Most naturalists would insist that the whole world is spatio-temporal and all the entities to be found in this world are studied by science. (Gasser, 2007, p. 4)

This suggests that the emphasis on spatiotemporal location is the result of a more fundamental understanding of the natural as that which can be investigated by scientific means. If so, we might be better off with a formulation like (5):

(5) Every entity is of a kind such that things of that kind can, in principle, be successfully investigated by science.

The 'successfully' is meant to avoid trivialization, since one could presumably attempt to investigate anything in a scientific fashion. Even given this, however, it's not obvious that (5) can rule out the unwelcome entities. If by 'science' we mean just any appropriately disciplined and rational enquiry, how much can (5) rule out? Certainly on that reading of 'science' it is not strong enough to rule out ghosts, entelechies, God, and other unwelcome entities; there is no reason to suppose such things could not be successfully investigated in a rational, disciplined fashion. Even if we restrict 'science' to empirical investigations, such entities are not obviously ruled out.

In light of this point, one might be tempted to the view that it is just a mistake to try to treat 'supernatural' or 'non-natural' as picking out a kind of entity. A striking thought experiment due to Ludwig Wittgenstein points in that direction:

[W]e all know what in ordinary life would be called a miracle. It obviously is simply an event the like of which we have never yet seen. Now suppose such an event happened. Take the case that one of you suddenly grew a lion's head and began to roar. Certainly that would be as extraordinary a thing as I can imagine. Now whenever we should have recovered from our surprise, what I would suggest would be to fetch a doctor and have the case scientifically investigated and if it were not for hurting him I would have him vivisected. And where would the miracle have got to? For it is clear that when we look at it in this way everything miraculous has disappeared. (Wittgenstein, 1965, p. 10)

Perhaps we should say that there is no difference between things that we see as 'miraculous' and those that are natural, that the felt difference lies entirely on the side of our attitudes. Being a naturalist might then amount simply to a readiness to investigate in a rational, disciplined fashion whatever phenomena one might encounter.<sup>7</sup> There would be no such thing as metaphysical naturalism.

If we give 'science' more content, however, (5) may yet yield a distinction between the natural and the unwelcome. A tempting thought here is that we should define metaphysical naturalism by reference specifically to the natural sciences—physics,

---

<sup>7</sup> Consider in this light the following remark from Bernhard Vollmer in a recent anthology of papers on naturalism: 'We understand naturalism as a natural-philosophical anthropological position. It can be characterized briefly by the thesis: everywhere in the world everything can be explained rationally' (Vollmer, 2007, p. 35).

chemistry, biology, geology, and so on. There are at least two different ways to pursue this thought. On one option, which parallels (5) closely, metaphysical naturalism becomes the claim that the techniques specific to the natural sciences are adequate for investigating the relevant kinds of things:

(6) Every entity is of a kind such that things of that kind can, in principle, be successfully investigated using the techniques specific to the natural sciences.

On a rather different option, the focus is not on the techniques specific to the natural sciences but on the substantive content of the theories developed in the natural sciences:

(7) Every entity can be fully described using the theories developed in the natural sciences.

One advantage of (7) is that it suggests a simple parallel with physicalism or materialism, where the latter doctrine might be understood simply as (7) with 'physics' replacing 'the natural sciences'.

Neither thesis is satisfactory as it stands. It is not obvious that there are techniques specific to the natural sciences; nor is it clear how the 'in principle' in (6) should be understood. With (7) the question arises as to just which theories developed in the natural sciences are at issue, and of course the 'fully described' clause leaves much to be desired as well. Questions of this sort are taken up in the literature on formulating physicalism, and parallel questions could be pursued for naturalism.

Depending on one's motivations for being a naturalist, (6) and (7) may seem objectionably arbitrary; those who count themselves as epistemic naturalists may insist there is no reason to focus on natural science in particular, so long as philosophy is deprived of any special authority. Such naturalists may accept the deliverances or techniques of the social sciences as providing warrant for philosophical claims even while denying (6) and (7).

### The 'nothing special' interpretation

Is there a way to make sense of the variety of different suggestions for how to understand naturalism? One promising avenue is what I call the 'nothing special' interpretation.<sup>8</sup>

The discussion of what might count as a 'natural entity' led us to a focus on the natural sciences; those sciences are 'natural' presumably because they study the world apart from humans. The domains proper to physics, chemistry, biology, geology, meteorology, astronomy, and so on are understood without reference to human activity or special human interests. These sciences deal with 'nature' in the specific sense of the world apart from humanity, while the others deal with humanity itself—those which, to use the older jargon, comprise the moral sciences. Metaphysical naturalism, I suggest, is best

---

<sup>8</sup> This interpretation is presented in Witmer, unpublished.

seen as the view that this division does not reflect any deep differences: humanity is nothing special, but just another part of nature.

The 'nothing special' interpretation can explain why naturalists are united in rejecting certain entities—e.g., ghosts, gods, and immaterial minds. One aspect of these examples is that such entities, if they exist, operate in a way quite foreign to the rest of the 'natural' world, at least as we understand that world to operate. But that is not on its own enough to account for their being intuitively supernatural; the discovery of a unique and unanticipated kind of energy would not thereby tempt us to declare naturalism falsified. However, another aspect of at least many of these examples is that the entities in question are akin to human beings in some relevant respect. In particular, they are often understood as agents endowed with free will. Crudely put: what makes ghosts objectionably supernatural is not just their being inexplicable in terms of the existing theories of natural science but, further, their being people, or at least agents. If we were to learn that ghosts exist, then we would conclude that there is a deep division between humanity, or at least human-like things, and the rest of nature.<sup>9</sup>

Not every example of an unwelcome entity is one of a human-like entity; recall Dupré's example of vital fluids and De Caro's example of 'prime movers unmoved'. Nonetheless, where the example is of an entity that is akin to a human being, we may say that it runs counter to naturalism because it requires a serious divide between things like us and the rest of the world. And even with these other examples, the entities plausibly are ones that are of special significance to humans. Consider entelechies—entities unlike anything one finds in the inorganic realm which play a crucial role in making life possible. We take special interest in living things, whether or not we find them especially similar to us. Such things as entelechies plausibly strike us as contrary to naturalism because of their combination of (i) being essentially involved in something of special interest to us (namely, life), and (ii) not being explicable in terms of the sciences that deal with those things that are not of special interest to us.

As a way of testing the 'nothing special' interpretation, we might manipulate some examples to see just when the threat to naturalism intuitively appears. Consider the earlier example of a newly discovered, unprecedented kind of energy; on its own, this seems compatible with naturalism. Now let us add that this energy is alleged to be directly controlled by the thought processes of appropriately situated humans. With that addition, we are much more likely to find the posit suspiciously 'supernatural'.

Can the 'nothing special' interpretation give us a clear metaphysical thesis? It suggests at least a schema. First, we delineate a number of phenomena that are of special interest to humans; call these the 'Human Interest Phenomena', or the HIP for short. (The acronym is deliberate; that which is 'hip' is the 'cool', the trendy, that which people celebrate and are interested in.) Second, there is the rest of the universe, that in which we have no special interest apart from our general intellectual curiosity; this is the non-

---

<sup>9</sup> An interesting case: suppose we discovered that ghosts exist, are psychologically much like us, but actual human beings are fully understandable in physical terms. Would naturalism be falsified? I expect many would say that it is still is. This is one reason I suggest a formulation in terms of 'human interest phenomena'—which would include things that are not themselves human but are still human-like.



HIP, which is akin to 'nature' as the HIP is akin to 'man', as potentially set apart from nature. Metaphysical naturalism might then be expressed thus:

(8) The HIP are not fundamentally different from the non-HIP.

At bottom, the idea is that there's nothing special about us, or about those things that are of special concern to us. This 'nothing special' idea is of course quite vague, and this vagueness is reflected in (8). 'Fundamentally different' admits of more than one reading, and the class of Human Interest Phenomena does not have sharp boundaries. Nonetheless, this seems to capture well the core metaphysical conviction of the naturalist, and in so far as we want a diagnosis of what people have in mind, such indeterminacy is only to be expected.

This core idea can be found in declarations of naturalism reaching back to the early 20th Century. In 1927, in clarifying his understanding of 'naturalism,' John Dewey credits his naturalism for finding incredible any view that posits a 'gulf ... between nature and man', adding that, on his view, 'human affairs ... are projections, continuations, complications, of the nature which exists in the physical and pre-human world' (Dewey, 1927, p. 58). In the present day we can find clear echoes of this basic idea. In a review of a recent anthology of papers on naturalism, Joseph Rouse suggests that one strand of naturalist thought is driven by an opposition to 'humanism':

Here lies the motivation for some naturalists' hostility to folk psychology, freedom, transcendental reason, the irreducibility of consciousness or first-person standpoints, and above all, any conception of normativity as sui generis. Human beings live in a world indifferent or even hostile to our interests, desires, values, or perspectival priorities, and the sciences provide our primary access to this anthropo-peripheral world to which we must accommodate ourselves. This anti-humanist strain of naturalism aspires to a hard-headed, resolute commitment to a thoroughly scientific self-understanding that can free us from the residual strands of self-aggrandizing illusion or wishful thinking that still confer disproportionate significance upon our all-too-human preoccupations. (Rouse, 2008)

It is, I think, fair to say that contemporary declarations of naturalism are often accompanied by a vivid awareness of just how tempting it is for us to endorse views of the world that are self-congratulatory in the way Rouse describes.

The 'nothing special' interpretation makes it plain just why, despite the variation in how exactly 'naturalism' is understood, we have near-unanimity in seeing naturalism as opposed to theism. After all, if theism is correct, the HIP are decidedly different from the non-HIP, since the lives of human beings are located at the very purpose of the universe, and the non-HIP merely provide a stage for the human drama. The present suggestion also makes it easy to see why the mind-body problem should loom so large in the concerns of naturalists without their conviction being limited to the status of the mind. There is no question that mentality is included in the HIP; whether we can see it as not fundamentally different from the rest of the world is thus a key question for the naturalist.

Another advantage of (8) is that it makes sense of the fact that some surprising metaphysical views might be seen as metaphysically naturalistic despite non-standard elements. Consider a version of theism according to which all of nature is equally sacred in the divine plan, where humans occupy no special place in the universe. Or consider a panpsychist view according to which the human mind is indeed of a piece with the non-HIP, as everything in the universe is endowed with mentality. In such cases, the HIP and non-HIP are fundamentally similar, but the similarity is achieved at the cost of a rather radical revision of how we understand the non-HIP, e.g. giving up the ordinary view that microphysical particles lack mental properties.

The deliberate vagueness of 'fundamental similarity' in (8) also enables us to see how a property dualist could (without radical revision of our understanding of the physical) take himself to be a naturalist. When, for example, David Chalmers describes the dualism defended in The Conscious Mind (1996) as naturalistic, he stresses the similarities between phenomenal properties and physical properties on the dualist view, which similarities are primarily a matter of being governed by laws of nature and being subject to similar causal explanations.<sup>10</sup>

The 'nothing special' interpretation is offered as a diagnosis more than as a recommendation, and it does not handle all cases of naturalist talk equally well. For instance, it is not as easy to see how it can make sense of the emphasis on spatiotemporal location we saw above with (4).<sup>11</sup> Two additional aspects of it are worth pointing out, however. First, it helps connect epistemic naturalism with metaphysical naturalism, since the key epistemic naturalist view might be described as the conviction that there is nothing special about what philosophers can do, and hence no position of special authority we can wield in adjudicating the development of science. Second, and most important for present purposes, it makes sense of the idea that physicalism is a version of metaphysical naturalism, as we'll see below.

### **Formulating physicalism**

Let us take the core idea of metaphysical naturalism to be (8). Presumably the physical realm is a decidedly non-HIP realm; there is nothing about physical entities as such that is of special interest to humanity. If we then formulate physicalism as the claim that everything is either physical or at least nothing over and above the physical, and that 'everything' ranges over all the HIP, then physicalism counts as a version of metaphysical naturalism. The significance of the physical, on this view, lies precisely in

---

<sup>10</sup> As he puts it, the view in question 'is naturalistic because it posits that everything is a consequence of a network of basic properties and laws' (Chalmers, 1996, p. 128).

<sup>11</sup> There might nonetheless be an intelligible link between the privileging of the natural sciences and the focus on spatiotemporal location we find in (4). Consider the following striking fact about the entities posited by the most admired of natural sciences, namely, physics: physical entities pervade all of space and time, including those regions we ourselves occupy. Taking this for granted, one might implicitly suppose that anything in contrast to the entities described by the natural sciences would have to be separate from this single spatiotemporal system. Being spatiotemporally located would not, on its own, confer the status of being natural, but the pervasiveness of entities already deemed natural conjoined with this claim could motivate one to think that anything that is fundamentally different from those natural entities would need to be located elsewhere, thus giving rise to (4) as an expression of naturalism.

its being a realm that is both indifferent to human interests and plausibly believed to be what ultimately makes up those things that are of special interest to us.

A formulation of physicalism needs to address at least three aspects of the claim: its extent, what is meant by 'physical', and the relation that everything in the extent is supposed to bear to the physical.

### The extent of physicalism

The question of the extent or scope of physicalism has received less attention than the other two questions of formulation. This may be due to the fact that controversies about the doctrine tend to revolve around cases that are obviously within its scope, such as actually existing mental phenomena. The preceding discussion of naturalism suggests that the physicalist claim ought to extend at least far enough to include everything that we might consider HIP—including biological phenomena, moral phenomena, and perhaps more.

What the physicalist claims about these phenomena will not be limited to the claim that the particular objects involved are physical objects. This is made clear by the fact that 'property dualism' as a position on the mind-body problem is considered incompatible with physicalism. The thesis should imply that all of the relevant properties and relations are either physical or—at least, when actually instantiated—appropriately nothing over and above the physical.

If properties are understood as abstract objects, this claim can strike one as very puzzling; after all, how can an abstract entity be a physical entity? Depending on how we want to understand 'physical', however, this may make perfectly good sense. For instance, classifying a property  $P$  as physical may mean only that it is a property in the domain of the relevant physical theory, where  $P$  itself can be construed along traditional Platonic lines as a necessarily existing entity not located in space or time. Most contemporary physicalists aim to be neutral on questions about the reality of abstract entities, and this stance is vindicated by the availability of such readings of 'physical'.

The focus on properties raises several other questions about the extent of physicalism. If we allow that there are uninstantiated properties, must they be included as well? Perhaps we can exclude those that are necessarily uninstantiated, but not those that are merely by luck uninstantiated. And there are nice questions about those that are necessarily instantiated—e.g. being such that  $2+2=4$ . One promising suggestion here is that the claim should extend just to those properties that are capable of being contingently instantiated.<sup>12</sup> Obviously, the significance of one's answers to these questions depends as well on just how many properties or relations one is willing to recognize in the first place, though the more permissive one is the more important it is to get clear on this parameter.

---

<sup>12</sup> A notable discussion of the extent of physicalism can be found in Andrew Melnyk's 2003 book [A Physicalist Manifesto](#), in which he proposes that the claim extend to all causal or contingent tokens (pp. 10-11). The 'causal' is not redundant; if a necessarily existing God existed but still has a causal impact, that token would also need to be caught in the physicalist net.

Further, we should ask whether an attention to properties is sufficient. The physicalist is perhaps obliged to make a further claim about objects, events, laws, facts, or other entities. Consider objects: since they are plausibly not just bundles of properties, a claim only about the properties of objects might not say enough to satisfy physicalist scruples. For another example, consider laws: depending on one's view of laws of nature, even if all the actual instances of properties are caught in the physicalist net, this may not be enough to rule out laws of nature that the physicalist should reject.<sup>13</sup>

Nonetheless, for the remainder of this article I will focus solely on actually instantiated properties (understood hereafter as including relations), as it is at least obvious that the physicalist will claim that those are physical or (at least on each actual occasion of instantiation) appropriately nothing over and above the physical.

### Specifying the 'physical'

Let us start by considering a distinction drawn in the middle of the last century in thinking about the mind-body problem. In a classic discussion of emergentism, Paul Meehl and Wilfrid Sellars (1956) defined 'physical<sub>1</sub>' and 'physical<sub>2</sub>' as follows:

Physical<sub>1</sub>: an event or entity is physical<sub>1</sub> if it belongs to the space-time network.

Physical<sub>2</sub>: an event or entity is physical<sub>2</sub> if is definable in terms of theoretical primitives adequate to describe completely the actual states though not necessarily the potentialities of the universe before the appearance of life. (Meehl and Sellars, 1956, p. 252)

The physical<sub>2</sub> might be more loosely described as the inorganic; Meehl and Sellars' qualification regarding the 'potentialities' of such things is meant to make room for the possibility that after the appearance of life new, unpredictable features may emerge. Those features would not need to be named in an account adequate to describing the pre-organic world, however, and hence would not count as physical<sub>2</sub>.

'Physical<sub>1</sub>' is reminiscent of an approach to the mind-body problem associated more with the term 'materialism' than 'physicalism,' namely, defining the fundamental category as matter, where matter is understood in the Cartesian fashion as that which is extended in space. It should also remind us of the formulation of naturalism as the thesis that everything is located in one space-time network. As argued above, that attempt to make sense of naturalism is problematic. The notion indicated with 'physical<sub>2</sub>' comes closer to being an adequate account of the natural, as the inorganic seems clearly to belong to the realm of things that fail to be of special human interest.

If the physicalist makes use of 'physical<sub>2</sub>', her doctrine threatens to be equivalent to metaphysical naturalism, whereas it seemed at first to be a logically stronger thesis. To restore that strength one could keep the approach exemplified in the definition of 'physical<sub>2</sub>' and narrow the domain of relevant explananda. In fact, if we think of physics as a discipline charged with explaining some range of phenomena less expansive than

---

<sup>13</sup> I have argued (Witmer, 2001) that the physicalist should include a claim about laws so as to ensure that physicalism is not a lucky accident, though I don't think we should make physicalism a necessary truth.

'the inorganic', we might define the physical by reference to physics, or, rather, to some idealized version of physics—a version that succeeds in explaining those target phenomena.

This is the route taken by what is likely the most popular contemporary approach for specifying the physical. Plausibly, the popularity of this strategy turns on the fact that physicalism is itself motivated in some way by the impressive success of physics. Insofar as what motivates the physicalist is the success of some broader science—say, the physical sciences, as opposed just to physics itself—then a definition in terms of this broader set of theories may be in order, including perhaps chemistry, astronomy, and more. At one extreme what may impress the physicalist might be nothing more than the 'natural sciences'—in which case we are driven back to metaphysical naturalism.

There are two reasons one might want to focus on physics in particular—or, more precisely, that part of physics that is fundamental by physicists' own lights, quantum mechanics and microphysical theory. First, the entities at issue in the other 'physical sciences' seem likely to be appropriately 'nothing over and above' the entities delineated by physics, so including them in the domain of the physical is unnecessary. One could use 'physical' to encompass both that which counts as physical by the criteria we are now searching for and that which is appropriately nothing over and above the physical according to that first criterion, but that is likely to court confusion. In this article I will use 'physical' to mean that which qualifies as physical according to the initial criterion, and will occasionally use 'properly physical' to stress this point when necessary.

The second reason one might want to limit the (properly) physical to that part of physics concerns the justification of physicalism. One of the most popular and promising ways of supporting the doctrine appeals specifically to the thesis that physics is causally complete—more precisely, the thesis that all physical events can be given entirely adequate causal explanations solely by appeal to other physical events and laws. The easiest way to justify that claim, in turn, is by understanding the properly physical in terms of the domain of quantum mechanics and microphysical theory (Sturgeon 1998).

If it is agreed that the physical should be limited in this way, then one must bear in mind the fact that much discussion of physicalism has been conducted as if the 'physical' includes a broader range of properties. For instance, J. J. C. Smart's celebrated defense of the identity theory (Smart, 1959) focused on the identity of mental properties with neurophysiological properties—which would not count as properly physical on the present suggestion. If such properties really are appropriately nothing over and above the physical as delimited by theory, then this is a harmless practice, but it may help avoid confusion to distinguish an ordinary notion of the physical on which we enjoy an intuitive grip from the official notion tied to theory.<sup>14</sup>

The proposal that we specify what 'physical' is to mean by referring to physical theory has prompted a considerable skeptical literature (Crane and Mellor, 1990; van Fraassen, 1996; Montero, 1999; Crook and Gillett, 2001; chapter 5 of Stoljar, 2010). The core

---

<sup>14</sup> Papineau appeals to 'paradigmatically physical' events (1993, p. 30); see also the discussion of the ordinary-physical as it relates to the theoretical-physical in Witmer, 2000.

worry can be traced to Carl Hempel (Hempel, 1969) and is now often referred to as 'Hempel's Dilemma'. It may be put as follows. The theory by reference to which the 'physical' is to be defined is either some actually existing theory or some theory that exists only in some ideal, possible future. In light of the history of science, it is likely that present theories are false, even if they are in some sense closer to the truth than past theories. As a result, the physicalist who takes the former option will be defining physicalism by reference to a false theory, which seems an unhappy result.<sup>15</sup> On the other hand, if the physicalist chooses the latter option, the resulting notion faces two important worries.<sup>16</sup> The obscurity objection is that an appeal to an ideal physical theory leaves the physicalist thesis too obscure to be properly evaluated. If we don't know what the ideal physical theory looks like, we don't know what sorts of properties will count as physical properties; how, then, can anyone support, attack, or otherwise assess the physicalist thesis? The inappropriate extension objection (to use the apt term from Wilson, 2006) is that the appeal to an ideal theory allows for the possibility that intuitively nonphysical entities could count as properly physical by being included in the ideal physical theory. Suppose that there are metaphysically fundamental mental events that make a causal difference to the behavior of physical particles, and an ideal physical theory will, because it must succeed in explaining the behavior of such particles, end up positing those mental entities. The result in that case is that such entities count trivially as physical, which seems an inappropriate extension of the category.

One important response to Hempel's Dilemma is that offered by Andrew Melnyk (Melnyk, 2003, pp. 223-237; also see Melnyk, 1997). Melnyk defends the first option (appealing to current physical theory) by arguing that physicalism is itself a scientific hypothesis, and hence, the attitude the physicalist takes towards it should be the same attitude that scientific realists take towards what they consider to be the best of current scientific theories. As scientific realists usually recognize that current theories are likely to be false, the attitude in question need not be one that is rendered irrational by that recognition. On Melnyk's account of the appropriate attitude, one assigns the theory in question a higher probability than any of its relevant rivals, where not every incompatible theory counts as a relevant rival. As a result, the physicalist can take this attitude towards current physical theory without thinking it true or even likely true. Even if we accept Melnyk's account of the relevant attitude, it's not clear that it resolves our dilemma. One reaction is to say that both the scientific realist and the would-be physicalist face a problem: shouldn't they be able to tell us what they do think is likely true?

Other responses to the dilemma can be classified as either blunting the second horn or avoiding both by understanding 'physical' without reference to theory at all. Among the

---

<sup>15</sup> It is worth noting that the mere falsity of current physical theory wouldn't matter; the physicalist is only exploiting the theory for its inventory of entities. The problem arises only insofar as one thinks that current physical theory makes an error in including or leaving out entities that an ideal theory would not.

<sup>16</sup> An additional worry is sometimes aired that is not, in my view, very serious. This is the triviality worry: if the ideal physical theory is ideal in the sense that it succeeds as a theory about all those phenomena which the physicalist believes to be ultimately physical, then those phenomena are trivially counted as physical simply by virtue of being the subject matter of the ideal physical theory. This worry can be set aside, however, as the physicalist can define the ideal theory as one that succeeds as a theory about all phenomena of interest to the physicist—which is less than all those believed by the physicalist to be ultimately physical. (This point is related to the one in the main text about exploiting Meehl and Sellars's definition of 'physical<sub>2</sub>'.)

former (Papineau, 1993; Wilson, 2006) is the strategy of dealing with the problem of inappropriate extension simply by stipulating that the 'physical' is not to include anything mental unless it is nothing over and above something else that qualifies as physical.<sup>17</sup> This move helps somewhat with the obscurity objection, since one can then know something about what is not included in the physical. An importantly different response to the dilemma is to define the physical in an entirely negative way. This may be called the *Via Negativa* response: what is meant by a physical property simply is a non-mental property—or, more precisely, one that can be understood in a way that doesn't characterize it in mental terms.<sup>18</sup> Another variant is to drop talk of the 'physical' altogether and simply identify a metaphysical thesis as, say non-mentalism: everything is either non-mental or nothing over and above the non-mental (Papineau and Spurrett, 1999). Both versions should be distinguished from the approach that combines an appeal to ideal physical theory and a negative stipulation, as that offers an indirect link to current physical theory, the impressiveness of which is again part of the motivation for many physicalists (for discussion, see Gillett and Witmer, 2001; Montero and Papineau, 2005).

#### The 'nothing over and above' claim

Suppose an account is available that tells us when a property counts as (properly) physical in the first place. As mentioned, many physicalists want to allow properties that are not themselves properly physical but are nonetheless related in the right way to the physical. Before focusing on how that 'right way' might be understood, it will prove instructive to get clear on why a simple identity approach is unpopular. Why not opt for (9)?

- (9) Every property is identical with a properly physical property.

Common lore has it that (9)—known also as the type identity thesis—was overthrown by the recognition that there are properties that can be multiply realized, that is, instantiated at different times by virtue of quite different physical features (Putnam, 1967; Fodor, 1974; Boyd, 1980). It is important to distinguish two aspects of multiple realizability. When it is said that a given property F is multiply realizable, one part of the claim is a denial that there is any physical property P such that F and P are necessarily coextensive. This claim is one that an opponent of physicalist might urge in attacking the doctrine; it can hardly motivate dropping (9) as a requirement for physicalism.

The other aspect of multiple realizability, however, is that in the cases under consideration, the relation between the physical and F is plainly one in which it is misguided to see the instantiation of F as something requiring more than what is already given in the relevant physical conditions. This is easiest to see (and historically was made most evident) when the property F is already understood as a functional property, that is, the property of having some property that plays such and such a role. Where the

---

<sup>17</sup> Similarly, if one wants to ensure that physicalism is a species of naturalism, one might expand this prohibition to include anything that is HIP—so that, for instance, there are no basic biological properties among the physical.

<sup>18</sup> The *Via Negativa* needs to be spelled out in a way that doesn't commit the physicalist to the falsity of the classical identity theory; we cannot characterize the view, then, by saying that on it physical properties are defined simply as those properties not identical with any mental properties. Some more elaborate characterization as that given in the main text is needed.

property that plays the role is physical, and its playing that role is due entirely to physical events and laws, then the instantiation of F comes along 'for free'; it is not genuinely additional to the physical conditions there satisfied. It is what Armstrong has called an 'ontological free lunch' (Armstrong, 1997, pp. 12-13).

The relation exemplified in these examples has come to be known as 'realization': the property F is realized by some physical property or properties in a way that satisfies physicalist requirements. Exactly how the term 'realize' is best used has been a matter of some contention (Gillett, 2002; Polger, 2007; Shoemaker, 2007); the present point, however, is just that the relevant examples illustrate how a property could fit into a physicalist world without being strictly identical with any physical property. Indeed, even if the examples are entirely fictional, so that no properties are ever actually multiply realized, if their possibility is allowed, that would be enough to show that (9) is not necessary for the truth of physicalism. This is an important point to keep in mind in light of the renewed controversy over whether there really are multiply realized properties (Kim, 1992; Heil, 1999; Shapiro, 2000; Polger, 2002).<sup>19</sup>

The foregoing discussion sheds some light on the relation between physicalism and 'reductionism'—at least in some sense of that vexed and politically loaded term. While 'reductionism' is often associated simply with the type identity thesis, the rationale for not requiring such identity we can, without that requirement, secure something that still has a 'reductive' sound to it, namely, the claim that nonphysical properties are nothing over and above the physical. It is far from obvious, however, just how this reductive commitment may or may not be related to claims about intertheoretic reduction as discussed in the philosophy of science (e.g. in such seminal texts as Kemeny and Oppenheim, 1956; Nagel, 1961; Schaffner, 1967; Nickles, 1973; Churchland, 1979; Hooker, 1981).<sup>20</sup>

Thus far, we might formulate physicalism as (10):

(10) Every property is either physical or, on every actual occasion of instantiation, nothing over and above the physical conditions that actually obtain.

One may wonder, of course, just how this 'nothing over and above' talk is to be understood. Such talk is certainly not without intuitive appeal, and it is associated with talk that has wide currency in philosophy generally, especially talk of grounding one property in another, or talk of something's having one property in virtue of having another. It is also associated with talk of truth-making, and the physicalist is plausibly committed to the claim that when a nonphysical property F is actually instantiated, the physical facts make true the proposition that F is thus instantiated (Melnyk, 2003, p. 59).

---

<sup>19</sup> It's worth pointing out, too, that even if functional properties are deemed impossible for some reason, there are other kinds of cases in which a property appears to 'come for free' given other, more basic properties, e.g. when a determinate is instantiated by a determinate.

<sup>20</sup> For a useful discussion of one example of how different literatures on 'reduction' can become confusingly entangled, see Robinson, 2001.



In moving beyond (10), we might aim to eliminate the 'nothing over and above' talk entirely in favor of more familiar notions, finding it objectionably obscure. Even if we don't think such elimination necessary to have a respectably articulated thesis on the table, we will likely want to clarify (10) by getting clear on just what it is meant to entail.

The notion of supervenience has been widely thought to help in this regard. Briefly, one family of properties supervenes on another family just in case there can be no difference in the former without some difference in the latter. In other words, the total state of something with regards to one family of properties (the 'subvenient' if you will) suffices for its total state with regards to the supervenient family. An easy example to use to keep the idea clear is that of aesthetic and non-aesthetic properties. Plausibly, if two items differ in their aesthetic properties, they must also differ in some non-aesthetic properties; there cannot be a bare aesthetic difference. This is in turn a function of the derivative status of aesthetic properties; if something is ugly, something makes it ugly.

If the nonphysical properties are an ontological free lunch, fixing the physical properties should fix the nonphysical properties; and this just what an appropriate supervenience thesis delivers. The qualifier 'appropriate' is crucial, however, since there are many different ways in which the basic idea of supervenience can be fleshed out. In particular, the relevant comparisons can be understood in a variety of ways. What sorts of objects are to be compared for sameness and difference in the relevant properties, and how exactly are the relevant pairs to be selected? We might compare objects, regions, or entire worlds, and we might even compare the state of one thing with respect to its supervenient properties with the state of a distinct thing to which it is mapped with respect to its subvenient-properties. Permutations of these parameters can result in a wide variety of distinct supervenience claims. (See Kim, 1993; Savellos and Yalçin, 1995; Stalnaker, 1996).

Perhaps the most popular supervenience thesis in contemporary work is that which makes use of the notion of a 'minimal physical duplicate,' a term introduced by Frank Jackson (Jackson, 1998, pp. 11-13) to enable us to formulate a kind of global supervenience thesis, one that compares entire possible worlds. A minimal physical duplicate of a world w is, intuitively, what one would get if one were to take the physical description of w as a recipe, building a new world out of nothing but the ingredients spelled out in that recipe. No additional objects, events, or properties are added; nothing is in that world except what comes along necessarily given the recipe. The supervenience thesis now available is (11):

(11) Any minimal physical duplicate of the actual world is a duplicate simpliciter of the actual world.

The need for minimality arises from an important complication. Consider the property of being the happiest person in the world. Even if mental properties are in fact as the physicalist envisages, it is clear that fixing all the physical properties will not on its own suffice for someone's having this property. To see this, consider a possible world w just like this one in every physical respect, yet in addition, w contains an extra mind that is entirely independent of the physical. Even if the mental state of each actual individual is as the physicalist imagines, if the extra individual in w enjoys a level of happiness

greater than that of each actual individual, this addition will make a difference as to who has the property of being the happiest, and, as a result w differs from the actual world with respect to some nonphysical property without differing in any physical respect.<sup>21</sup>

In light of this complication, it may be incorrect to describe the physicalist's commitment as requiring that the way things are generally is nothing over and above the way things are physically; after all, to get something that suffices for the distribution of nonphysical properties one needs not just the way things are physically but that in combination with a kind of totality fact—the fact that nothing is added to the situation beyond that which comes for free with the physical properties. This totality fact is presumably not best classified as just a fact about the properly physical.

Suppose (as many hold) that (11) is necessary for physicalism; is it also sufficient? It will clearly not be sufficient if the range of possible worlds at issue in (11) is implicitly restricted in some way—say, to only those worlds with the same laws of nature as the actual world. In that case, the truth of (11) could be explained in a way incompatible with physicalism. There is no *prima facie* difficulty in supposing there could be laws of nature that link distinct entities in such a way as to guarantee that fixing the physical fixes everything else, even though the 'everything else' includes things that are indeed something over and above the physical. For (11) to come close to a sufficient condition, the relevant minimal physical duplicates must be drawn from among all the possible worlds, without restriction.<sup>22</sup>

Even with this clarification, there are reasons to suspect that (11) is not sufficient. For one, we might worry that there could be 'brute' metaphysical necessities, that is, necessities not explainable in terms of anything deeper, so that physical entities are simply related to nonphysical entities in a way that makes (11) true without having any consequences for the nature of those nonphysical entities (Melnyk, 2003, pp. 58-59). For another, the contrast between metaphysical and nomic necessity presupposed in the previous paragraph might be a false one, in which case even the 'full strength' reading of (11) might be explained by laws of nature, which apparently undermines its significance (see Wilson, 2005). Finally, there is a general lesson from the philosophy of science that seems applicable. The physicalist will want to say that the nonphysical properties are instantiated in virtue of the physical properties, which is an explanatory claim. But the familiar lesson from philosophy of science is that sufficiency is not itself sufficient for explanation; we may suspect, then, that a merely modal thesis such as (11) cannot ensure the relevant explanatory import of physicalism. In general, with each of these worries, the point is that there may be ways of explaining the truth of (11) that run counter to the physicalist thesis. What is needed is not just supervenience but, to use Terence Horgan's terminology, superdupervenience: supervenience that is explainable in a physicalistically acceptable fashion (Kim 1990; Horgan, 1993; McLaughlin, 1995).

---

<sup>21</sup> This example presumes that physicalism can be contingently true, so that even if all actual mental properties are nothing over and above the physical, there are possible instances that conform to the dualist's view of things. Functionalists have generally allowed such a possibility, but there has been some dispute over the availability of this position (Levine and Trogon, 2009).

<sup>22</sup> It should be noted that the minimal physical duplicates plausibly need to duplicate the laws of physics in the actual world, since such laws would be included in the physical 'recipe'; the point is that no other laws, if there are any other laws, can be presumed to hold constant.

In light of this, we might try formulating physicalism in terms of some set of possible relations that actually instantiated properties might bear to the physical, each of which would make available the kind of explanation of the sufficiency of physical facts that is wanted, an explanation that would vindicate the 'nothing over and above' talk. Consider (12):

(12) Every actually instantiated property is either (i) identical with some properly physical property or (ii) a functional property, where on each actual occasion of instantiation, it is realized by physical properties.

The idea that physicalism is best formulated as (12) has gained in popularity in recent years, largely because it seems better placed to provide a sufficient condition on the doctrine (Melnik, 2003; Kim, 1998, pp. 23-27).

Advocates of (12) will naturally focus on the notion of realization, given its central role and the fact that, unlike the clear and simple relation of identity, clarification is surely in order. On some accounts of realization, only properties that are functional in a suitably broad sense can be realized, where a property is functional in the sense used earlier, namely, being a property expressible as the property of having some property that plays role R (for a careful overview of some options here, see Shoemaker, 1981). A property R realizes such a functional property just in case it in fact meets condition C. This is a very simple and tidy account, but it exacts a considerable cost: any actually instantiated property that is not plausibly identified with a properly physical one needs to be functional in the relevant way, and it is far from obvious so many properties could be plausibly construed as functional in this way.

Not all accounts of realization tie the notion to functional properties understood in this fashion, however. Some alternative accounts lean on a certain increasingly popular background view, namely, that properties in general are individuated by the causal powers they contribute to their bearers. Suppose that a property F is associated with a set C(F) of causal powers in this way and that R is likewise associated with a set C(R) of causal powers. On the 'subset view' of realization, R realizes F just in case C(F) is a proper subset of C(R) (Clapp, 2001; Shoemaker, 2001; Shoemaker, 2007). Since C(R) includes members not in C(F), R is a distinct property from F, but since every member of C(F) is a member of C(R), once R is instantiated, F is thereby automatically instantiated as well.

An aspect of the subset view of realization that makes it especially appealing for the physicalist is that it seems to promise an account of how a property could be distinct from its physical realizer yet connected to it in a way that appears to avoid duplicating causal powers. When F is realized by R, it is not as if there are two sets of causal powers at work; there is just one set that is operative in the situation, while the two distinct properties are related in different ways to that one set of causal powers. This might alleviate several worries about the causal role of properties that are not themselves properly physical, although the matter is quite controversial.<sup>23</sup>

---

<sup>23</sup> The worries in question revolve around the 'exclusion problem'—roughly, the problem of finding a suitable causal role for nonphysical properties given the (presumed) fact that physical properties already provide adequate causal explanations on their

## Support for naturalism and/or physicalism

The doctrines of naturalism and physicalism enjoy extremely widespread support in contemporary philosophy, so much so that they seem often taken for granted. Nonetheless, there are at least two important ways in which physicalism has been supported, and there is a rough (and rarely articulated) argument for the more general and nebulous doctrine of metaphysical naturalism. I start with the latter.

### Naturalism and a general trend of dethronement

How can the thesis of metaphysical naturalism be supported? If the 'nothing special' interpretation is adopted, it is not hard to see how someone could find metaphysical naturalism compelling. Again, let 'HIP' stand for the Human Interest Phenomena—minds, life, morality, and perhaps other things—and consider (8) as an expression of metaphysical naturalism:

- (8) The HIP are not fundamentally different from the non-HIP.

It has often been said that there were a number of powerful historical moments in leading humanity to disabuse itself of the idea that it was at the center of the universe: there was the Copernican revolution, so that our planet is literally not at the center of the solar system; there was the Darwinian revolution, which assimilated mankind to the rest of the kingdom of life; and there was the impact of Freud, forcing us to take seriously the effects of our non-rational nature. These three famous moments seem to exemplify a more general trend of dethronement of humanity in the development of the sciences. It is not clear that one could specify this trend with any precision; it is nonetheless quite plausible to say that such a trend has shaped the thinking of contemporary philosophers—not to mention contemporary thought in general. The idea is simple and familiar: science has repeatedly debunked those views on which humanity and our interests occupy a special place in the world.

Whether this idea can be worked up into a powerful argument for metaphysical naturalism as per (8) is a good question. We could stipulate a more precise meaning for 'not fundamentally different', taking as our cue the formulation of the 'nothing over and above' aspect of physicalism. The more difficult task lies in thinking about how one could appropriately project from examples of the HIP that have been shown to be nothing over and above the non-HIP to the general conclusion that the same is true of all the HIP. Suppose that the class of the HIP is akin to the class of grue things (Goodman, 1983), so that generalizing from the naturalistic status of some of them to all of them is unwarranted. Similarly, the category of the non-HIP may be heterogeneous as well, undermining such a generalization.

A distinct strategy here would be to develop an argument aimed not at supporting (8) directly but instead at the conclusion that we should not give any epistemic weight to any inclination we may have to treating the HIP as special—or, perhaps, more weakly,

---

own. For representative discussions, see Heil and Mele, 1993; Horgan, 2001; Marcus, 2001; Heckmann and Walter, 2003; Bennett, 2007.

that we should treat them as *prima facie* unjustified, given their poor historical track record. On its own, this does not motivate endorsing (8); general considerations of simplicity may, however, tilt the scales here. If our inclinations to deny (8) are not to be trusted, and in general we should aim for simpler theories, then this may give us good reason to treat the truth of (8) as at least a default view.

Of course, if we have independent evidence for physicalism, and physicalism implies (8), then we have much more reason to endorse metaphysical naturalism.

### Physicalism: inductive arguments

The first of the two influential routes to justifying physicalism is roughly inductive in character: argue that since a wide variety of phenomena have already been shown to be nothing over and above the physical, the physicalist is justified in making the inductive leap to conclude that all actual phenomena are likewise grounded wholly in the physical.

One way to develop this argument is by appealing to an alleged history of successful intertheoretic reductions. That is certainly the justification imagined by those philosophers writing in the middle of the twentieth century who were concerned to defend a physicalist theory of the mind (see especially Oppenheim and Putnam, 1958). As they saw it, the work had already been done in showing that the sciences—at least, most of those other than psychology and those dependent on psychology—reduce, in order, to biology, chemistry, and ultimately to physics. In effect, on this view, physicalism as a view about everything other than the mental had already been demonstrated by a string of successful reductions, so that all that remained was to justify the extension of physicalism to the realm of the mental.

There are two important problems facing this approach. First, it is not entirely clear how intertheoretic reduction, understood as a phenomenon studied in the philosophy of science, is to be related to the key metaphysical claims of the physicalist. The literature on the former is bound up with epistemic concerns about the preservation of epistemic virtues across theory change, which is not directly relevant to the physicalist's claim. In particular, the best known account of intertheoretic reduction—the derivational account offered by Ernest Nagel (1961)—is subject to important worries regarding its significance for metaphysical debate (Kim, 1998).

Second, and most pressing, the view that intertheoretic reductions are widespread in actual science is now widely rejected by philosophers of science. The claim that biology has been shown to be reducible to chemistry, for example, is far from being uncontroversial. No one doubts that all sorts of interesting intertheoretic links have been found and developed, but the controversy is over the proper interpretation of those achievements (See chapter 2 of Garfinkel, 1981; Kitcher, 1984; Smith, 1992; Sklar, 1999; for a book-length treatment of the significance of 'post-reductionist philosophy of science', see Horst, 2007.)<sup>24</sup>

---

<sup>24</sup> It is a nice question whether or not the shift in attitude between the middle of the last century and now is due more to changes in how carefully philosophers of science attend to the details of actual science or more to shifting standards of what is needed for something to count as a genuine 'intertheoretic reduction'. The answer to this question is relevant to the prospects for the argument strategy discussed in the next paragraph.

Even if the physicalist cannot find a history of successful intertheoretic reductions, she may still find a history in which physicalist commitments are repeatedly vindicated. One very detailed development of this strategy can be found in the last chapter of Melnyk's A Physicalist Manifesto (2003), in which he reviews a variety of examples from natural science (ranging from such things as chemical bonding to cardiac musculature), to argue that the scientific results give us the resources for defending many identity hypotheses of the sort that vindicate physicalism. On his view, physicalism requires every particular entity to be either physical or physically realized, where realization is tied to the notion of a functional type. The scientific results will not speak directly to this, of course, since they won't speak of 'realization' directly or focus on these questions; but it may be that, as Melnyk argues, the results do enough to show that if we make the needed identifications, we are able to explain all that needs explaining about the relevant phenomena.

The general strategy may be illustrated with what Melnyk says about solidity. As he stresses, 'the physicist's account of solids just sketched is not yet a physicalist account of solids' (Melnyk, 2003, p. 241). The former implies that solidity includes certain properly physical features, which features can explain a wide variety of explananda—e.g., the fact 'that solid items retain their size and shape despite the effects of gravity, and despite the application to them of (a restricted range of) mechanical forces' (242). Still, this is consistent with saying that solidity involves something in addition to what is needed to explain those things. The key move is then made by pointing out that, so far as those explananda are concerned, 'there is no theoretical need to construe solidity as anything more than a physical or a functional-but-physically-realized property' (242) Given the general principle that we should minimize our commitments, this gives us reason to think that the best explanation of those facts is one that vindicates the physicalist hypothesis regarding solidity. If this can be repeated for many and varied cases, this can serve as powerful evidence for physicalism in general.

The task is far from trivial. One potential difficulty turns on the fact that the relevant explananda might be described in the first instance in terms far removed from the properly physical, so that even if the physicalist hypothesis can adequately explain the physical aspects of target explananda, it may fail to explain all that needs explaining. Further, there are important questions about how inference to the best explanation is to work when supporting identity claims in general (Kim, 2005), and those may force some rethinking of this strategy.<sup>25</sup>

### Physicalism: causal arguments

---

<sup>25</sup> A somewhat different way one might use the history of science to argue for physicalism is via what one can call the argument from proven methodological utility. Instead of focusing on the content of scientific results, this argument would turn on the claim that scientific practice has often presupposed that physicalism is true, and this presupposition is at least partly responsible for its success. If the history of science shows that it pays to presume physicalism, we could argue that the best explanation of this is that physicalism is true. One important difficulty with this strategy arises simply from the general difficulty in adjudicating claims about presuppositions. Even if it is plausible to say that actual scientists have presumed physicalism, more work is needed to show that this presumption—and not, perhaps, some less weighty or significant presumption—played an important role in shaping the theories constructed, experiments designed, or the like (Gillett, 2001).

A rather different kind of approach to supporting physicalism makes use of a general thesis regarding the physical, namely, the causal completeness of physics (also often called the 'causal closure' of physics). The idea, informally put, is that one never needs to advert to anything other than physical events and the laws of physics to provide an adequate causal explanation of physical events. Physics thus provides a domain which is complete in that it is never lacking for the resources to provide causal explanations of events in its own domain.

The completeness thesis does not say that the only legitimate causal explanations of physical events are ones that appeal exclusively to physical factors; its claim is only that, for every physical event which can be causally explained at all, there exists a causal explanation of it that appeals only to physical conditions and laws and is entirely adequate on its own, so that if that were the only explanation to be had, there would be nothing unsatisfactory about that explanation. As such, it is the sort of thesis that can be supported by considerations internal to physical theory, by judging the success of physical theory in producing such explanations, without having to rely on any claims about the relation of physical theory to anything else. This is a considerable advantage, as any such claims are likely to be philosophically contentious.

If the causal completeness of physics is granted, the physicalist may press the following point. Consider some property M that is not, at least not obviously, a properly physical property. If, as seems likely, M is involved in the causation of some physical event, then, unless we say that M is either identical with some physical property or appropriately nothing over and above such properties, we face a puzzling situation. If e is the physical event at issue, we know from the causal completeness of physics that it can be given an adequate explanation adverting only to other physical conditions; since M is also involved in causing e, its role would seem as a result to be redundant—a consequence that seems at least odd, and intolerable if we were forced to accept it as a widespread occurrence. If the only way to avoid this unhappy consequence is to accept that M conforms to the physicalist thesis, then we have what seems a tidy and apparently powerful argument for physicalism (Papineau, 1993; Sturgeon, 1998).<sup>26</sup>

The argument is not as tidy as it may first appear, however. A major complication arises from the fact that causation appears to be a relation between events, yet the above argument attempts to draw conclusions about the nature of the property M, which I described as being 'involved' in the causation of a physical event. If we stick simply with events, we can perhaps draw the conclusion that every event is either identical with a physical event or nothing over and above some physical event. As noted earlier, however, the physicalist commitment clearly extends to properties. So it is crucial that the causal argument be developed with careful attention to the role of properties in

---

<sup>26</sup> The causal argument should be sharply distinguished from the traditional attack on Cartesian dualism that questions the intelligibility of causation spanning the mental and physical divide. The argument does not concern the mental/physical divide in particular, nor does it turn on claims about intelligibility at all. It should also be distinguished from Davidson's famous argument (Davidson, 1970) for the thesis that every event is a physical event. While there are a few (indirect) relations between this argument and Davidson's, the latter turns fundamentally on the claim that there are no strict laws that can be given in non-physical terms, whereas the present argument could happily allow such laws. For an excellent discussion of these different arguments and their independence from each other, see chapter 6 of Foster, 1991.

causation. It is easy enough to say that a property may have 'causal relevance' in a given cause-effect pair just in case the cause event was a cause in virtue of instantiating that property; but how should this complication be incorporated into the argument for physicalism?

Another considerable difficulty concerns the relationship between avoiding the conclusion that the putatively nonphysical property is causally redundant and establishing that it is either properly physical or nothing over and above the physical. The argument presumes that the only way to avoid an offensive kind of overdetermination is to embrace physicalism about the relevant property, at least as it is then and there instantiated. But it's not obvious just how to establish this. There is again a considerable literature related to this issue, though its relation is somewhat indirect. Much of the literature is devoted to what Kim has made well known as the 'exclusion problem': this is the problem for the would-be 'nonreductive' physicalist who wants to maintain the causal relevance of those properties that he nonetheless counts as not reducing to the physical (Heil and Mele, 1993; Kim, 1998; Horgan, 2001; Bennett, 2007). It is not always clear just how 'reductive' and 'nonreductive' are being used in this discussion, though if we take 'reductive' to mean a commitment to a simple type identity hypothesis, we can see the dispute as revolving around the following question: How can anything less than strict identity avoid the threat of problematic overdetermination? Answers to this question will be indirectly relevant to the more general question: exactly what are the possible ways in which a problematic kind of overdetermination can be avoided? Addressing that question, in turn, is crucial for assessing the causal argument for physicalism.

## References

Armstrong, D. M. (1978), Universals and Scientific Realism 1: Nominalism and Realism. Cambridge: Cambridge University Press.

Armstrong, D. M. (1997), A World of States of Affairs. New York: Cambridge University Press.

Bennett, K. (2007), 'Mental Causation.' Philosophy Compass 2, (2), 316-337.

Boyd, R. (1980). 'Materialism Without Reductionism: What Physicalism Does not Entail'. in N. Block, (ed), Readings in the Philosophy of Psychology. Cambridge, MA: Harvard University Press, 67-106.

Chalmers, D. J. (1996), The Conscious Mind: In Search of a Fundamental Theory. Oxford: Oxford University Press.

Churchland, P. M. (1979), Scientific Realism and the Plasticity of Mind. Cambridge: Cambridge University Press.

Clapp, L. (2001), 'Disjunctive Properties: Multiple Realizations'. Journal of Philosophy, 98, (3), 111-136.



Crane, T. and D. H. Mellor (1990), 'There is No Question of Physicalism'. Mind, 99, (394), 185-206.

Crook, S. and C. Gillett (2001), 'Why Physics Alone Cannot Define the "Physical": Materialism, Metaphysics and the Formulation of Physicalism'. Canadian Journal of Philosophy, 31, (3), 333-360.

Davidson, D. (1970), 'Mental Events', in L. Foster and J. W. Swanson, (eds), Experience and Theory. Amherst, University of Massachusetts Press, pp. 79-101.

De Caro, Mario. (2010), 'Varieties of Naturalism', in R. C. Koons and G. Bealer (eds), The Waning of Materialism. New York: Oxford University Press, pp. 365-374.

Dewey, J. (1927), 'Half-Hearted Naturalism'. Journal of Philosophy, 24, (3), 57-64.

Dupré, J. (2004), 'The Miracle of Monism', in M. DeCaro and D. Macarthur (eds), Naturalism in Question. Cambridge, MA: Harvard University Press, pp. 36-58.

Fodor, J. (1974), 'Special Sciences, or the Disunity of Science as a Working Hypothesis'. Synthese, 28, 97-115.

Foster, J. (1991), The Immaterial Self. New York: Routledge.

Garfinkel, A. (1981), Forms of Explanation. New Haven, Yale University Press..

Gasser, G., ed. (2007), How Successful Is Naturalism? Frankfurt: Ontos.

Gillett, C. (2001), 'The Methodological Role of Physicalism: A Minimal Skepticism', in Gillett and Loewer (2001), pp. 225-250.

Gillett, C. (2002). 'The Dimensions of Realization: A Critique of the Standard View'. Analysis, 64, (4), 316-323.

Gillett, C. and B. Loewer, eds. (2001). Physicalism and Its Discontents. Cambridge: Cambridge University Press.

Gillett, C. and D. G. Witmer (2001), 'A "Physical" Need: Physicalism and the Via Negativa'. Analysis, 61, (4), 302-309.

Goldman, A. I. (1999), 'A Priori Warrant and Naturalistic Epistemology'. Philosophical Perspectives, 13, 1-28.

Goodman, N. (1983), Fact, Fiction and Forecast. Cambridge, MA: Harvard University Press.

Heckmann, H-D. and S. Walter, eds. (2003), Physicalism and Mental Causation: The Metaphysics of Mind and Action. Charlottesville: Imprint Academic.

Heil, J. (1999), 'Multiple Realization'. American Philosophical Quarterly, 36, (3), 189-208.

Heil, J. and A. Mele, eds. (1993), Mental Causation. New York: Clarendon Press.

Hempel, C. G. (1969), 'Reduction: Ontological and Linguistic Facets', in S. Morgenbesser, P. Suppes and M. White, eds., Philosophy, Science and Method: Essays in Honour of Ernest Nagel. New York, St. Martin's Press, pp. 179-199.

Hooker, C. A. (1981), 'Towards a General Theory of Reduction - Parts 1-3'. Dialogue: Canadian Philosophical Reviews, 20, 38-59, 201-236, 496-529.

Horgan, T. (1993), 'From Supervenience to Superdupervenience: Meeting the Demands of a Material World'. Mind, 102, (408), 555-586.

Horgan, T. (2001), 'Causal Compatibilism and the Exclusion Problem'. Theoria, 16, (40), 95-116.

Horst, S. (2007), Beyond Reduction: Philosophy of Mind and Post-Reductionist Philosophy of Science. New York: Oxford University Press.

Jackson, F. (1998), From Metaphysics to Ethics: A Defence of Conceptual Analysis. New York: Oxford University Press.

Kemeny, J. G and P. Oppenheim. (1956), 'On Reduction'. Philosophical Studies, 7, 6-17.

Kim, J. (1990), 'Supervenience as a Philosophical Concept'. Metaphilosophy, 21, (1-2), 1-27.

Kim, J. (1992), 'Multiple Realization and the Metaphysics of Reduction'. Philosophy and Phenomenological Research, 52, 1-16.

Kim, J. (1993), Supervenience and Mind: Selected Philosophical Essays. Cambridge: Cambridge University Press.

Kim, J. (1998), Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation. Cambridge, MA: MIT Press.

Kim, J. (2003), 'The American Origins of Philosophical Naturalism', in R. Audi (ed), Philosophy in America at the Turn of the Century. Charlottesville, Philosophy Documentation Center, pp. 83-98.

Kim, J. (2005), Physicalism, or Something Near Enough. Princeton: Princeton University Press.

Kitcher, P. (1984), '1953 and All That: A Tale of Two Sciences'. Philosophical Review, 93, 335-374.

Ladyman, J. and D. Ross (2007), Every Thing Must Go: Metaphysics Naturalized. New York: Oxford University Press.

Levine, J. and K. Trogon (2009), 'The Modal Status of Materialism'. Philosophical Studies, 145, (3), 351-362.

Lewis, D. (1983), 'New Work for a Theory of Universals'. Australasian Journal of Philosophy, 61, 343-377.

Maddy, P. (2001), 'Naturalism: Friends and Foes'. Philosophical Perspectives, 15, 37-67.

Marcus, E. (2001), 'Mental Causation: Unnaturalized but not Unnatural'. Philosophy and Phenomenological Research, 63, (1), 57-83.

McLaughlin, B. (1995), 'Varieties of Supervenience', in Savellos and Yalçin (1995), pp. 16-59.

Meehl, P. E. and W. Sellars. (1956), 'The Concept of Emergence', in H. Feigl and M. Scriven (eds), The Foundations of Science and the Concepts of Psychology and Psychoanalysis. Minneapolis: University of Minnesota Press, pp. 239-252.

Melnyk, A. (1997), 'How to Keep the "Physical" in Physicalism'. Journal of Philosophy, 94, (12), 622-637.

Melnyk, A. (2003), A Physicalist Manifesto: Thoroughly Modern Materialism. Cambridge: Cambridge University Press.

Montero, B. (1999), 'The Body Problem'. Nous, 33, (2), 183-200.

Montero, B. and D. Papineau (2005), 'A Defence of the via negativa Argument for Physicalism'. Analysis, 65, (287), 233-237.

Nagel, E. (1961), The Structure of Science. New York: Harcourt Brace World.

Nickles, T. (1973), 'Two Concepts of Intertheoretic Reduction'. Journal of Philosophy, 70, 181-201.

Oppenheim, P. and H. Putnam (1958), 'Unity of Science as a Working Hypothesis', in H. Feigl, M. Scriven and G. Maxwell (eds), Minnesota Studies in the Philosophy of Science Vol. 2. Minneapolis: University of Minnesota Press, pp. 3-36.

Papineau, D. (1993), Philosophical Naturalism. Oxford: Blackwell.

Papineau, D. and D. Spurrett (1999). 'A Note on the Completeness of "Physics"'. Analysis, 59, (1), 25-29.

Polger, T. W. (2002). 'Putnam's Intuition'. Philosophical Studies, 109, (2), 143-170.

Polger, T. W. (2007), 'Realization and the Metaphysics of Mind'. Australasian Journal of Philosophy, 85, (2), 233-259.

Putnam, H. (1967), 'Psychological Predicates' (later retitled 'The Nature of Mental States'), in W. H. Capitan and D. D. Merrill (eds), Art, Mind and Religion. University of Pittsburgh Press, pp. 37-48.

Quine, W. V. O. (1969), 'Epistemology Naturalized'. Ontological Relativity and Other Essays. New York: Columbia University Press, pp. 69-90.

Robinson, H. (2001), 'Davidson and Nonreductive Materialism: A Tale of Two Cultures', in Gillett and Loewer (2001), pp. 129-151.

Rouse, J. (2008), Review of Gasser 2007, in Notre Dame Philosophical Reviews. <http://ndpr.nd.edu/review.cfm?id=12284>.

Savellos, E. E. and Ü. D. Yalçın, eds. (1995), Supervenience: New Essays. New York: Cambridge University Press.

Schaffner, K. F. (1967), 'Approaches to Reduction'. Philosophy of Science, 34, 137-147.

Sellars, R. W. (1927), 'Why Naturalism and Not Materialism?' The Philosophical Review, 36, (3), 216-225.

Shapiro, L. A. (2000), 'Multiple Realizations'. Journal of Philosophy, 97, (12), 635-654.

Shoemaker, S. (1981), 'Some Varieties of Functionalism'. Philosophical Topics, 12, (1), 83-118.

Shoemaker, S. (2001), 'Realization and Mental Causation', in Gillett and Loewer (2001), pp. 74-98.

Shoemaker, S. (2007), Physical Realization. New York: Oxford University Press.

Sklar, L. (1999), 'The Reduction (?) of Thermodynamics to Statistical Mechanics'. Philosophical Studies, 95, (1-2), 187-202.s

Smart, J. J. C. (1959), 'Sensations and Brain Processes'. Philosophical Review, 68, 141-156.

Smith, P. (1992), 'Modest Reductions and the Unity of Science', in D. Charles and K. Lennon (eds), Reduction, Explanation, and Realism. New York: Oxford University Press, pp.19-43.

Stalnaker, R. (1996), 'Varieties of Supervenience'. Philosophical Perspectives, 10, 221-224.

Stoljar, D. (2010), Physicalism. New York, Routledge.

Stroud, B. (2004), 'The Charm of Naturalism', in M. DeCaro and D. Macarthur (eds), Naturalism in Question. Cambridge, MA: Harvard University Press, pp. 21-35.

Sturgeon, S. (1998), 'Physicalism and Overdetermination'. Mind, 107, 411-432.

Van Fraassen, B. C. (1996), 'Science, Materialism, and False Consciousness', in J. Kvanvig (ed), Warrant in Contemporary Epistemology: Essays in Honor of Plantinga's Theory of Knowledge. Lanham: Rowman and Littlefield, pp. 149-181.

Vollmer, G. (2007), 'Can Everything Be Rationally Explained Everywhere in the World? Theses and Declarations for Naturalism', in Gasser (2007), pp. 25-48.

Wilson, J. (2005), 'Supervenience-Based Formulations of Physicalism', Nous, 39, (3), 426-459.

Wilson, J. (2006), 'On Characterizing the Physical'. Philosophical Studies, 131, (1), 61-99.

Witmer, D. G. (2000), 'Locating the Overdetermination Problem'. British Journal for the Philosophy of Science, 51, (2), 273-286.

Witmer, D. G. (2001), 'Sufficiency Claims and Physicalism: A Formulation', in Gillett and Loewer (2001), pp. 57-73.

Witmer, D. G. (2006), 'How to Be a (Sort of) A Priori Physicalist'. Philosophical Studies, 131, (1), 185-225.

Witmer, D. G. (Unpublished), 'Making Sense of "Naturalism"'.

Wittgenstein, L. (1965). 'A Lecture on Ethics'. Philosophical Review, 74, 3-12.