On the Distinction between Reductive and Non-Reductive Physicalism
Forthcoming in Metaphilosophy (July 2011)
Matthew C. Haug
Department of Philosophy
College of William & Mary
P.O. Box 8795
Williamsburg, VA 23187
mchaug@wm.edu

Abstract

In this paper, I argue that the debate between reductive and non-reductive physicalists is best characterized as a disagreement about which properties are natural. Among other things, natural properties are those that characterize the world completely. All physicalists accept the “completeness of physics,” but I argue that this claim contains a subtle ambiguity, which results in two conceptions of natural properties. I argue that reductive physicalists should assert, while non-reductive physicalists should deny, that a single set of low-level physical properties is natural in both of these senses. This way of drawing the distinction succeeds where previous approaches have failed and illuminates why the debate about reductionism is important.

Keywords: physicalism, reduction, non-reductive physicalism, completeness of physics, natural properties

1. Introduction

In a recent paper, Karen Bennett (2008) claims that non-reductive physicalists have sometimes lost track of what they are committed to. According to Bennett, non-reductive physicalists should not claim that mental properties are not identical to any physical properties. Rather, they should just insist that mental properties are not identical to narrow- or restricted-sense physical properties.
Bennett’s claim makes use of the widely accepted fact that there are several notions of ‘physical’ at work in the literature. First, there are *restrictive* definitions of ‘physical’ like “pertaining to or connected with matter; material; opposed to psychical, mental, spiritual” (Oxford English Dictionary, second edn., 1989). Opposed to these are *inclusive* understandings of ‘physical,’ which are almost synonymous with “natural,” such as: “concerned with natural laws and forces or material things” or “of, pertaining or relating to, or in accordance with, the regular processes or laws of nature” (ibid.).

Contemporary accounts of the restricted sense of ‘physical’ typically contrast “low-level physical” (Chalmers 1996, 33) properties with a disjoint set of “high-level physical” or “ordinary physical” properties. “Low-level physical properties,” according to Chalmers, are “the fundamental properties that are invoked by a completed theory of physics. Perhaps these will include mass, charge, spatiotemporal position; properties characterizing the distribution of various spatiotemporal fields, the exertion of various forces, and the form of various waves; and so on” (ibid.).¹ High-level and ordinary physical properties and events include *arm movements*, *earthquakes*, *being a toaster*, and *having e-fibers firing in a human nervous system*. The inclusive-sense physical domain is thus the union of the low-level (restricted-sense) and high-level (broad-sense) physical domains. Importantly, if non-reductive physicalism is true, then mental properties are inclusive-sense physical (since they are broad-sense physical). This is what distinguishes non-reductive physicalism from dualism.

While I think that Bennett is correct that non-reductive physicalists can, and should, claim that mental properties are identical to broad-sense physical properties (merely in virtue of

¹ Note that properties like *mass* and *charge* are also possessed by macro-physical objects. Thus, they are not felicitously characterized as “microphysical” as Chalmers does. For this reason, I prefer to use the phrase ‘restricted-sense physical’ instead of ‘microphysical’ even though the latter is more evocative.
being self-identical), I believe that the reductive physicalist can legitimately ask why the non-reductive physicalist is entitled to claim that mental properties are broad-sense physical (but not restricted-sense physical) and what this claim amounts to. What unifies these two senses of physicality? Why isn’t the non-reductive physicalists identity claim just a mealy-mouthed way of endorsing a form of dualism?

This is apparently a basic point underlying some of Jaegwon Kim’s attacks on non-reductive physicalism. Kim would take Bennett’s view to be a form of reductive physicalism (as she notes) because he believes that there is no stable position that is distinct from both reductive physicalism and dualism. According to Bennett, whether non-reductive physicalism is a genuine competitor to reductive physicalism, dualism, and emergentism depends on whether there are well-defined, physicality-preserving “methods of construction” (such as mereological fusion and other forms of composition) applied to restricted-sense physical properties that the non-reductive physicalist accepts but the reductive physicalist rejects.

Without examining the details of these methods of construction, it is not clear that much progress has been made in spelling out, much less resolving, what exactly is at issue between reductive and non-reductive physicalists. Further, it is unclear how a difference of opinion about which “methods of construction” are legitimate has implications for the metaphysical, semantic, and epistemic issues that Bennett herself sees as the interesting questions about reduction (e.g. whether mental properties “survive” multiple realization, whether our mental terms refer to “second-order” properties or their realizers, and issues about classification and individuation of properties). Finally, on this account of the difference between reductive and non-reductive physicalism, it is difficult to see how these various issues hang together (if at all). Do they all arise out of a core disagreement among physicalists?
If Bennett is correct to think that non-reductive physicalism cannot be captured by simple slogans such as “mental properties are not identical to physical properties,” how can the distinction between reductive and non-reductive physicalism best be captured? As the debates about mental causation, causal exclusion, and multiple realizability have become more intricate, we risk losing track of why we care whether reductive or non-reductive physicalism is true. Although it is helpful to point out that simple slogans often obscure hidden subtleties, one should be careful not to get lost in those subtleties, lest the contrast between reductive and non-reductive physicalism appear to be a distinction without a difference. Further, although the resources of analytic metaphysics and philosophy of language have clarified and furthered these debates, it is important to remind ourselves of a point Jerry Fodor made over thirty years ago. Although the debate between reductive and non-reductive physicalism may be at bottom an empirical one, it is a debate that has broad normative implications for the methodology of science and for the proper way of understanding the mind’s place in the natural world (Fodor 1974, 97). Any attempt to articulate the differences between reductive and non-reductive physicalism should preserve this insight. It should make it clear why the contrast is important – why it matters if reductive or non-reductive physicalism is true.

In this paper, I defend an account of the contrast between reductive and non-reductive physicalism that addresses these issues. According to this account, the difference reflects a deep disagreement about natural properties. Among other things, natural properties are supposed to give a complete characterization of the world. As I argue in Section 4, a subtle ambiguity in the general claim that “physics is complete” gives rise to a distinction between two conceptions of natural properties. I claim that this distinction gets to the heart of the disagreement between reductive and non-reductive physicalists. Roughly, reductive physicalists should claim that the
single set of restricted-sense physical properties satisfies both conceptions of naturalness, while non-reductive physicalists should deny this. I show how framing the debate in this way illuminates the importance of debates about reduction. Before developing the positive proposal, in Sections 2 and 3 I argue for some prima facie constraints that a successful attempt at drawing the distinction between reductive and non-reductive physicalism should meet and argue that prominent existing proposals fail to do so.

2. Prima Facie Constraints on an Adequate Characterization of the Distinction

There is arguably no monolithic doctrine known as “non-reductive physicalism.” Rather, debates between reductive and non-reductive physicalists are complex and multifaceted, and non-reductive physicalists hold subtly different views about a range of related issues concerning causation, explanation, and the nature of events and properties. However, there seem to be a couple of core theses concerning the explanatory autonomy of psychology and other special sciences and the inadequacy of appealing solely to restricted-sense physical properties in a complete account of the natural world, which the majority of non-reductive physicalists endorse. Of course, what exactly autonomy amounts to and what is required for a complete account of the world are partially what is at issue in the debate. Characterizing the debate between reductive and non-reductive physicalism should not prejudice which view is correct and should provide the means to spell out what constitutes explanatory autonomy.

Any attempt to provide an account of the core difference between reductive and non-reductive physicalism – what I will call “drawing the distinction” – should satisfy a fairly weak principle of charity: one should attempt to avoid drawing the distinction in a way that implies that either side of the debate has committed an obvious error, is radically mistaken about the nature of her own position, or is defending an obviously inconsistent view.
This principle of charity underwrites some plausible prima facie constraints on any adequate attempt to draw the distinction. It should be largely faithful to the broad contours of existing literature: Jaegwon Kim and David Lewis should turn out to be reductive physicalists, while Jerry Fodor and Hilary Putnam should be counted as non-reductive physicalists. Of course, if this paper is correct, a successful attempt at drawing the distinction will not be entirely faithful to the literature. For, as I argue in the next section, existing accounts of the distinction are unsatisfactory. Further, an adequate proposal for drawing the distinction should be neutral: it should not by itself imply that reductive physicalism is true and non-reductive physicalism is false (or vice versa). In the same vein, all other things being equal, one attempt at drawing the distinction is superior to another if it does not result in non-reductive physicalism turning out to be rhetorical or notational variant of either (a) reductive physicalism or (b) dualism and strong forms of emergentism (according to which configurations of physical entities result in the emergence of novel fundamental forces). Of course, in the end, Kim may be right that non-reductive physicalism is an unstable or incoherent view. However, this should not be implied merely by an attempt to say what is at issue in the debate.

Now, it may turn out that some aspects of the debate between reductive and non-reductive physicalism hinge on subtle details of the correct account of causation. Perhaps Kim’s presentation of the causal exclusion argument depends upon the assumption that low-level physical events do “all the causal work,” and perhaps the coherence of this assumption in turn depends on a flawed view of causation. However, even if this were so, it would show only that one particular argument against non-reductive physicalism is flawed but would presumably not undermine the entire debate.
I should emphasize that these are merely *prima facie* constraints on drawing the distinction. In the end, the entire debate may turn out to be ill-conceived or ill-motivated, but charity recommends that we endorse this conclusion only after all other interpretations are exhausted. It would be implausible if some of the best philosophers of mind in the last forty years were *completely* mistaken in thinking that there is a substantive debate here that is not identical to debates about, say, the nature of events or causation.

3. Previous Attempts at Drawing the Distinction

3.1. Token-Identity vs. Type-Identity

The standard way of drawing the distinction between reductive and non-reductive physicalism is to claim that both reductive and non-reductive physicalists agree that every token event (or property instance) is identical to a physical event (property instance), but that non-reductive physicalists deny, while reductive physicalists assert, that every event type (or property) is identical to a physical event type (property) – in short, that non-reductive physicalism can be characterized as endorsing token-identity while denying type-identity (see Fodor (1974, 100); Davidson (1980)).

I claim that this is not a successful means of drawing the distinction. Characterizing non-reductive physicalism in this way is uncharitable in that it fails to be neutral, by failing to distinguish between reductive and non-reductive physicalism, while ensuring that they are both versions of physicalism.

The literature in this area is confused and confusing because authors are often talking past one another regarding the nature of events (see note 3 for an example). Theories of events fall into two broad classes: (i) the view that events are property exemplifications and (ii) the view

---

2 Of course, some non-reductive physicalists reject token identity as well. I think the main task for such philosophers is to explain how their view remains a physicalist one.
that events are concrete particulars (regions of space-time or the contents of those regions). On either account, token-identity conjoined with the denial of type-identity is not an adequate characterization of non-reductive physicalism.

Consider the former view of events. Kim has long defended the view that events are exemplifications by objects of properties at a time (Kim 1976). On this conception of events, an event is a “structured complex” and can be represented by an ordered triple, \( \langle x, P, t \rangle \), where \( x \) is the constitutive object (n-tuple of objects) of the event, \( P \) is the constitutive property (n-adic relation) or “generic event,” and \( t \) is the time when \( x \) has \( P \) (objects in the n-tuple bear the n-adic relation to one another). Given this conception, a property instance is a token event – \( x \)’s having \( P \) at \( t \).³

³ One example of authors talking past one another: Louise Antony assumes that what Kim means by “property instances” are the entities that have the property “as opposed to tropes” (Antony 1999, 43 n.3). But this is a flawed interpretation of Kim’s argument and ontology. Interpreted as Antony does, the upshot of Kim’s argument for the identity of property instances would be the unremarkable fact that system \( s \) is identical to itself – not a claim that one would think that Kim would need to argue for – since, as Kim makes clear, he thinks that substances (objects) have properties. One might be drawn to Antony’s interpretation of Kim since it is presumably objects/substances that have causal powers, not tropes or “particularized properties.” However, this also betrays a misunderstanding of Kim’s ontology. For, he thinks that events (property instantiations/exemplifications) have causal powers (Kim 1999, 16). Antony’s interpretation would make more sense if one adopted the view that events (i.e. concrete particulars) have properties. But this is not Kim’s view, events do not have properties; they are partly “constituted” by properties, or they exemplify properties. So Antony is not agreeing with Kim but rather talking past him. Of course, adopting Antony’s ontology, both non-reductive and reductive physicalists will agree that every mental “property instance” is identical to some physical “property instance” or other. However, this amounts to claiming that the same subjects have mental and physical properties. This claim, coupled with the non-reductionist’s denial of type-identity, makes it unclear how non-reductive physicalism differs from dualism. So the neutrality constraint is not satisfied.
instance\(^4\) implies that property instances of multiply realized properties are identical to the instances of the properties that realize them on a given occasion. Suppose that \(s\) is a system and \(E\) is a property realized by properties \(Q_1, Q_2, \ldots\). Then, according to Kim:

\[
s'\text{'}s \text{ having } E \text{ on this occasion is identical with its having } Q \text{ on this occasion.}
\]

There is no fact of the matter about \(s'\text{'}s \text{ having } E \text{ on this occasion over and above } s'\text{'}s \text{ having } Q\). Each instance of \(E\), therefore, is an instance of one of \(E\)’s realizers, and all instances of \(E\) can be partitioned into \(Q_1\)-instances, \(Q_2\)-instances, \(\ldots\), where the \(Q\)’s are \(E\)’s realizers. Hence, the \(E\)-instances reduce to the \(Q_i\)-instances. (Kim 1999, 15-6)

On this view of events, it is contradictory to assert property instance (event token) identity together with property (event type) distinctness. For, the claim that every token mental event is identical to a token physical event amounts to the claim that every triple, \(\langle x, M, t \rangle\), is identical to a triple \(\langle x, P, t \rangle\), where \(M\) is a mental property and \(P\) is a physical property. But, given the individuation condition on events, this implies that the constitutive properties are identical, that the mental property \(M\), is identical to the physical property \(P\). If one adopts this view of events, token-identity implies type-identity. (As [name removed for blind review] reminded me, Kim is aware of this problem and in response suggests that mental properties are not “constitutive properties of events” (Kim 1993, 364-5 n.5). I think this is best captured by the idea that mental properties are not natural. See Section 4 below.) So, formulating non-reductive physicalism as including the commitment to token identity is not neutral; the ontological claims of non-reductive physicalism collapse into those of reductive physicalism. To avoid this, the non-reductive physicalist must reject type-identity for properties (or events). Consequently, she is forced to reject token-identity for properties (events) as well, if she adopts Kim’s view of events. Without spelling out just how token mental events are related to physical events, non-

\(^4\) In some formulations of the causal inheritance principle, Kim admits that the causal powers of the realized property may be a proper subset of those of the realizer. Kim’s argument for identity does not go through with this weaker principle.
reductive physicalism now collapses into dualism or emergentism, and the neutrality constraint is again violated.

The main competing view of events, associated with W.V.O. Quine and Donald Davidson’s early views, sees them as concrete particulars – regions of space-time (or the contents of such regions). However, the claim of mere token-event identity, in this framework, is not stronger than substance physicalism. On the Davidsonian view of events, one can compatibly assert token-identity while rejecting type-identity. But, asserting token event identity now amounts to claiming that the same region of space-time (or its contents) has two properties, a mental one and a physical one. There is no requirement that the mental property even be supervenient or dependent on the physical property. Consequently, the neutrality constraint is not satisfied. It is not clear that this characterization of non-reductive physicalism differs from reductive physicalism, on the one hand, or mere substance physicalism, which itself is compatible with non-physicalist forms of property dualism and emergentism, on the other.\(^5\)

For convenience and definiteness, I assume a Kim-type account of events in this paper. However, given this framework, any non-reductive physicalist must deny both token- and type-identity for events. Thus, more must be said about the relation between mental and physical properties in order to distinguish non-reductive physicalism from dualism and emergentism.

How can the non-reductive physicalist deny both token- and type-identity claims for events, while remaining a physicalist? Does claiming that mental properties supervene on physical properties define a position which is both distinct from emergentism and other competitors to physicalism and which does not collapse into reductive physicalism? The answer

---

\(^5\) For a more detailed argument, on different grounds, that token physicalism is not a coherent, substantive view that is stronger than minimal substance physicalism and weaker than property (type) physicalism see Latham 2003.
to this question plausibly depends on the strength and nature of the supervenience relation (i.e., is it strong or weak, global or local, involving metaphysical or nomological necessity), on which there is an immense literature. The dominant view seems to be that strong (local or global) supervenience involving metaphysical necessity is required for physicalism.

However, it is controversial whether even this is enough to guarantee that a supervenience-based formulation of physicalism is distinct from its non-physicalist competitors. For instance, Jessica Wilson (2005) has argued that supervenience-based formulations of physicalism do not provide a genuine alternative to emergentism if one adopts a “necessitarian” view of the laws of nature (see also Melnyk (2003)). If Wilson is right, any form of supervenience-based “physicalism” is compatible with emergentism, and hence part (b) of the neutrality constraint is violated. Of course, even if Wilson is correct, there may still be some distinction between supervenience-based reductive and non-reductive “physicalism.” However, it would not be the distinction that has been lying behind the decades-long debates about the autonomy of the special sciences (since, as noted above, both versions would be compatible with emergentism). So, the overarching goal of illuminating why the debate between reductionism and non-reductionism matters would also not be met.

This suggests a deeper problem with using relations between mental and physical properties to attempt to differentiate reductive and non-reductive physicalism. Relations like supervenience, realization, and determination (or Bennett’s gesture at “methods of construction” discussed above) are technical notions that can be defined in a variety of ways (as witnessed by the huge body of literature on supervenience and the ever-growing number of articles on realization). One could attempt to use these relations to try to draw the distinction, but this just raises the further question of how to distinguish reductive from non-reductive versions of these
relations. For instance, Andrew Melnyk (2003) has argued that realization-based physicalism preserves all of the claims that a reductionist should care about, whereas other proponents of realization take it to be part of a non-reductive view. No progress has been made in characterizing the debate; the question merely reappears in a different context.

3.2. Functional Properties

As is well-known, there are two main varieties of functionalism in the philosophy of mind: those, like Sydney Shoemaker’s, which claim that there are functional roles in addition to the realizers of those roles and identify mental properties with the former, and those, like David Lewis’s, which deny that functional roles correspond to properties distinct from their realizers. Perhaps this contrast provides the means for drawing the distinction. Reductive physicalism would amount to the claim that no properties are functional (i.e. no properties possess their causal-functional roles essentially), while non-reductive physicalism would claim that some properties are functional (or, more contentiously, that all properties are, if one adopted the causal theory of properties).

I think that this is a more promising attempt to draw the distinction. However, the debate about which, if any, properties are functional is clearly not concerned with abundant properties (see below). For example, proponents of the causal theory of properties claim merely that a privileged subset of the abundant properties (the set of natural ones) is functional. Thus, this proposal in effect presupposes a distinction between natural and unnatural properties (see Section 4). It has struck on the right kind of contrast but has not gotten to the theoretical heart of the distinction.
4. The Distinction Between Reductive and Non-reductive Physicalism

All physicalists are committed to the hypothesis that all of the entities in the empirical world do not belong to distinct ontological domains. However, reductive and non-reductive physicalists interpret this hypothesis in very different ways. As I argue below, reductive physicalists should claim that all natural properties are restricted-sense physical. There are no natural special science properties (and hence no special science causal processes) but merely special science concepts or property designators – different ways of referring to restricted-sense physical properties. As domains of natural properties, the restricted-sense and inclusive-sense physical domains are held to be coextensive. The non-reductive physicalist should deny this. She should hold that there are natural physical properties that are not restricted-sense physical. The set of natural inclusive-sense physical properties cannot be identified with the set of natural restricted-sense physical properties.

4.1. Natural Properties

So what does it mean to say that a property (event, state, or process) is natural? Properties in one sense are easy to come by. Assuming an ontology that includes merely possible objects, any set of possible objects, or any function from possible worlds to extensions, will correspond to a property. According to this view, properties are abundant. For example, the set containing the golden mountain, Bob Dylan, my refrigerator, and the Atlantic Ocean is such a property. Grueness is a property in this sense, as is the property of being an incar (Hirsch 1980).

Clearly, abundant properties will not be appropriate for many of the roles that properties are called upon to serve, for example, accounting for similarity between objects and contributing causal powers. To arrive at something to serve these roles, one might propose singling out a special subset of the abundant properties, the natural ones. Many claims have been made
regarding what distinguishes natural properties from unnatural ones. For instance, Lewis writes the following about natural properties: “Sharing of them makes for qualitative similarity, they carve at the joints, they are intrinsic, they are highly specific, the sets of their instances are *ipso facto* not entirely miscellaneous, there are only just enough of them to characterise things completely and without redundancy. […] What physics has undertaken, whether or not ours is a world where the undertaking will succeed, is an inventory of the [natural] properties of this-worldly things.” (1986, 60). “Natural properties would be the ones whose sharing makes for resemblance, and the ones relevant to causal powers” (1983, 13).

These passages suggest several criteria that a property must meet in order for it to be natural:

*Resemblance:* only natural properties underwrite similarity between objects.

*Causal powers:* only natural properties contribute causal powers.

*Completeness:* the set of natural properties provides the minimal basis with which to characterize the world completely and without redundancy.

---

6 I will set aside the complication that naturalness comes in degrees, according to Lewis, as well as the problem of how to define degrees of naturalness. With this added complication, the distinction between non-reductive and reductive physicalism amounts to a dispute about whether or not non-restricted-sense physical properties are ever *sufficiently* natural.

7 Jonathan Schaffer (2004) identifies three similar “qualifications for the office of [natural] property”: similarity, causality, and minimality, the last of which is the claim that “[natural] properties serve as a minimal ontological base.” Schaffer argues that the minimality constraint should be abandoned and replaced with a *primacy* role; where the primary/derivative contrast corresponds to the difference between the ontological structure of reality (what is primarily real) and the linguistic truths (what is derivative or projected). If this is done, Schaffer argues that properties drawn from “all levels of nature” are primarily real from the start, and thus that the scientific properties are the only natural ones because they best fill *all* of the requisite roles. I think that this move is analogous to defining physical properties as those that are needed in an ideal theory of the world, i.e. one that runs afoul of the “vacuity” or “no determinate content” horn of Hempel’s dilemma. It threatens to make the distinction between natural and unnatural properties vacuous. For example, does this view hold that *being grue* is natural? Or that *being an incar* or *being a klable* are? If we ignore considerations of redundancy, as Schaffer (2004, 100) claims we should, aren’t they on the ontological side, not the linguistic side, just as much as neurons and beliefs are? It seems that they are and that Schaffer’s emendation lets too many
However, as I show, there are two senses of completeness – two things that can be meant by “characterizing things completely and without redundancy”: causal sufficiency completeness and fundamental category completeness.\(^8\)

Most philosophers who write about completeness hold that it is only the basic, microphysical domain that is complete, by which they apparently mean that it is only this microphysical domain that provides a minimal, complete domain (e.g., Baker 1993, 79; Sturgeon 1998, 415). (After all, it is hard to see how simply adding superfluous macroscopic physical entities could make the formerly complete microphysical domain incomplete.) One common formulation of completeness runs as follows:

\[ \text{Causal sufficiency completeness of physics: Every physical event is completely causally determined, insofar as it is causally determined, by restricted-sense physical causes.}\(^9\) \]

(I have intentionally left open how the first occurrence of ‘physical’ should be interpreted. Obviously, a stronger claim results if we read ‘physical’ in an inclusive sense, including not only microphysical events but also macroscopic, everyday and special science physical events, like my arm’s moving, a cell’s respiring, and a hurricane’s developing.)

We can, of course, abstract away from the physical domain and think of this kind of completeness as a schema for any domain of properties, X:

\[ \text{Causal sufficiency completeness: Every event of type X is completely causally determined, insofar as it is causally determined, by causes of type X.}\]

---

Properties count as natural. As discussed in the text, I do not think that the minimality/completeness role should be replaced; rather, it needs to be disambiguated.\(^8\) The remainder of this section draws heavily from Haug 2009.\(^9\) The ‘insofar as…’ proviso is normally included to account for the possibility of quantum indeterminacy. Depending on one’s philosophical predilections, completeness can be given an epistemic formulation, in terms of concepts, theories, or explanations, or it can be formulated metaphysically, in terms of properties, powers, causes, or causal sequences. In this paper, I remain as neutral as possible between these formulations.
One argument for the claim that only microphysics is complete appeals to the idea of levels in nature (or in our explanatory schemes). The thought is that only the categories at the fundamental, microphysical level will provide for exceptionless generalizations stated in a single, unified vocabulary. As Scott Sturgeon puts it:

Whereas biology admits that biological effects sometimes have non-biological causes, and psychology admits that psychological effects sometimes have non-psychological causes—and, more generally, special science S admits that S-effects sometimes have non-S causes—physics does not admit that physical effects have non-physical causes. Rather, physics considers itself closed and complete. It says physical effects have their chances fully determined by physical events alone. (1998, 413; cf. Baker (1993, 79))

The general idea of “descending to a lower microlevel” in order to explain anomalous phenomena at a higher level is common in the literature. As Kim has remarked (crediting the physicist David Bohm): “each time we descend to a lower microlevel, we do so because the current level is not causally closed (‘explanatorily complete’ may be a better term in this context)” (2005, 67). Kim’s parenthetical remark suggests that, in addition to causal sufficiency completeness, the general claim that physics is complete involves another claim that is not solely concerned with causation—namely, that we do not need to appeal to non-physical categories to provide a complete account, not only of the individual sufficient causes in the world, but also of the fundamental categories which, among other things, compose those causes. The categories that are used in special science generalizations are not sufficiently general to provide for a complete account of the fundamental nature of all the events that causally affect the physical universe. It is only a privileged set of physical categories that can accomplish this task. This completeness claim can be put as follows:

*Fundamental category completeness of microphysics*: microphysics provides a complete inventory of the fundamental kinds and categories (e.g. force, field,
space-time structure, mass/energy, particle) that are sufficient for a metaphysical basis for all of the events that causally impinge on the microphysical universe.\textsuperscript{10}

This way in which physics is thought to be complete is suggested by David Papineau’s claim that: “Current physics, I take it, aims to develop a complete theory of paradigm physical effects in terms of the categories of energy, field and spacetime structure” (1993, 31). Although Papineau does not mention this, the term “categories” here does not refer to the sufficient causes (i.e. events) of physical effects themselves but to the “building blocks” of those causes. That is, this empirical claim is not about sufficient causes but about fundamental forces (if forces are identified with fields) and other constituents of events. To deny this kind of completeness of physics, Papineau claims “would in effect postulate an extra mental force alongside the fundamental physical forces of gravity, the electroweak force, and the strong nuclear force” (1996, 4). But another way to deny it would be to claim that there is some fundamental category that was non-physical.\textsuperscript{11}

Taking this kind of completeness as a schema yields:

\textit{Fundamental category completeness:} kinds and categories of type X provide a complete inventory of the fundamental kinds and categories that are sufficient for a metaphysical basis for all of the events that causally impinge on events of type X.

To see that \textit{fundamental category completeness} is a distinct claim from \textit{causal sufficiency completeness} (using the physical domain for definiteness), first note the simple fact that the former completeness claim concerns general kinds, like quark, mass/energy, and electromagnetic

\textsuperscript{10}A metaphysical basis is a set of categories that characterizes the causal features and (if this is something additional) the intrinsic nature of a given set of phenomena.

\textsuperscript{11}One of Terence Horgan’s formulations of completeness also suggests the idea expressed by \textit{fundamental category completeness}. He claims that the causal completeness of physics “means that non-physical properties cannot be causally basic properties—ones that generate fundamental forces that combine with physical forces to yield net forces different from the net resultants of physical forces” (1993, 573).
field. These general kinds are not the sort of thing that can appear in individual causal sequences or be used to causally explain individual physical effects, which are what is at issue in causal sufficiency completeness.

Further, even if we consider instances of these general kinds, the accounts that are promised by fundamental category completeness are not entirely causal as is required by causal sufficiency completeness. Rather, these explanations will be, in part, constitutive ones in which the existence and properties of a whole are explained by the existence and properties of its parts. Perhaps this constitutive explanation will be a relatively deflationary one, involving the claim that constitution is identity. Or perhaps it will deny that constitution is identity and utilize a realization or supervenience relation. In any case, this synchronic relation that grounds the explanation of the properties and powers of a whole in terms of the properties and powers of its parts is the relation in virtue of which special science entities are “nothing over and above” fundamental physical entities. This relation is distinct from the diachronic, causal one that underwrites the explanations promised by causal-explanatory completeness.

Given that these are two distinct senses of completeness, the claim that natural properties “characterize things completely and without redundancy” could be interpreted as a claim about causal sufficiency completeness or as a claim about fundamental category completeness. There are potentially two sets of natural properties: natural$_{FC}$ properties, the smallest set of which fundamental category completeness is true and, natural$_{CS}$ properties, the smallest set of which causal sufficiency completeness is true. In the remainder of the paper, I use this ambiguity to articulate the core difference between reductive and non-reductive physicalists.

4.2. Natural Properties and Drawing the Distinction
Some of David Lewis’s and Jaegwon Kim’s comments about broadly physical properties like pain and heat suggest that differences of opinion regarding natural properties might be relevant to the distinction between reductive and non-reductive physicalism. For example, Lewis claims that the property of being in pain “cannot occupy ... any ... causal role because it is excessively disjunctive, and therefore no events are essentially havings of it. To admit it as causally efficacious would lead to absurd double-counting of causes” (1994, 307). Lewis also claims that the property (of space-time regions) containing rapidly moving particles is “a fairly natural, intrinsic property” while the property containing whatever phenomenon occupies the heat-role “is highly disjunctive and extrinsic” (1983, 45). So, Lewis claims, events involving the former are genuine (viz. natural) and causally efficacious, while events involving the latter are “too unnatural” and thus “inefficacious in the sense that [they cannot] figure in the conditions of occurrence of the events that cause things” (1983, 44-5).

Kim uses a similar line of reasoning to reach the same conclusion: “Given that mental kinds are realized by diverse physical causal kinds, therefore, it follows that mental kinds are not causal kinds, and hence are disqualified as proper scientific kinds. Each mental kind is sundered into as many kinds as there are physical realization bases for it, and psychology as a science with disciplinary unity turns out to be an impossible project” (1992, 327). According to Kim, the heterogeneous and disjunctive nature of multiply realized mental kinds means that they are unnatural, unscientific and not causal.

So I am suggesting that the distinctively reductionist feature of (particular temporal stages of) Lewis’s and Kim’s philosophy of mind is not the (species-restricted) type-identity claims that they make about broadly physical properties and events, but rather the claim that broadly physical properties and events are not natural. This claim should be common ground
amongst all reductionists. This way of drawing the distinction holds that the core dispute between non-reductive and reductive physicalism is about which properties are natural and, ultimately, about the more basic question of what *naturalness* of properties is – particularly, whether one set of properties can fulfill all of the roles that natural properties have been called upon to play.

Consistent reductive physicalists (those that are not tempted by non-reductionist intuitions) claim that there is a single set of natural properties – that the set of restricted-sense physical properties is the smallest set of which both *fundamental category completeness* and *causal sufficiency completeness* are true. Indeed, Lewis sometimes uses “fundamental” and “perfectly natural” interchangeably (e.g., 1994, 291).\(^\text{12}\)

On the other hand, according to non-reductive physicalism, inclusive-sense physical properties, including mental properties, are *natural* since on this view the set of inclusive-sense physical properties is the smallest set of which *causal sufficiency completeness* is true (along with fulfilling the other roles of natural properties). However, non-reductive physicalists should agree with reductive physicalists that only restricted-sense physical properties are *natural*. I think that this latter claim is needed in order to ensure that their theory is a substantive version of physicalism and is distinguishable from (not a terminological variant of) emergentism or

\(^\text{12}\) Reductive physicalists may want to count properties that are sufficiently “close” to the fundamental ones as natural. For instance, in debates about the mind-body problem it is often assumed that chemical and physiological properties are natural physical properties, even though they neither appear in basic physical theory nor are definable in basic physical terms. Whether they can consistently do this is an important question and is related to the question of whether the “exclusion problem” generalizes to all special science properties. A related point: at the end of his 1998 book, Kim outlines a view according to which functional properties and micro-based (structural) properties, like *being water*, count as physical (112ff.). He claims that we must do so in order for physics to be causally complete. Whether or not this is a concession to non-reductive physicalism depends in part on whether such properties can be multiply realized or multiply based and still count as natural. (see, e.g., Block 2003).
dualism. That is, we need the fact that restricted-sense physical properties provide a complete inventory of all fundamental forces in order to guarantee that the powers contributed by broadly physical properties do not require any novel fundamental forces or causal powers.¹³

Further, both reductive and non-reductive versions of physicalism agree that the set of inclusive-sense physical properties supervenes (with metaphysical necessity) on the set of restricted-sense physical properties. If the distribution of restricted-sense physical properties is fixed, then this fixes the distribution of the inclusive-sense properties. This of course implies that the distribution of causal powers is also fixed by the distribution of restricted-sense physical properties. But it does not imply that these restricted-sense physical property instantiations individually contribute each and every causal power. If a non-reductive physicalist rejects this claim, this amounts to rejecting the causal inheritance principle (see Pereboom 2002 for an argument that such a rejection is consistent with physicalism).

However, even if a non-reductive physicalist accepts causal inheritance, their view need not collapse into reductive physicalism. It is common ground in the debate that at least in worlds with the same laws, properties are individuated with respect to the causal powers that they, at least typically, contribute. Given this, when we ask about the smallest set of properties of which causal sufficiency completeness is true, the restricted-sense physical properties may not serve (even if configurations of them contribute every causal power) because they do not group these causal powers so as to pick out the smallest cluster of powers that was responsible for a certain high-level physical effect. That is, accepting causal inheritance does not imply that the broad-sense physical properties are unneeded for a complete inventory of minimal natural groupings of causal powers.

¹³ Proposals like Schaffer’s (see note 7) that simply abandon the completeness or minimality constraint are consequently not clearly physicalist theories. For a related view, see Baker 1993.
How does my proposal fare with respect to the prima facie constraints outlined above? It clearly satisfies the *neutrality* constraint. Reductive and non-reductive physicalism turn out to be distinct versions of physicalism that are not trivially true or false. Showing definitively that the *faithfulness* constraint is met would require (at least) another paper, since it would involve an extensive survey of the large literature on physicalism. However, two lines of evidence from this literature support the claim that a dispute about natural properties lies at the center of the debate. First, as mentioned above, the writings of Kim and Lewis explicitly highlight the centrality of naturalness. Second, claims about natural properties are arguably at the heart of multiple realizability arguments for non-reductive physicalism. The alleged fact that a special science property is multiply realizable by low-level physical properties is used to support the claim that it is not identical to any low-level physical property. Assuming that only low-level physical properties are candidates for natural\textsubscript{FC} properties and that some special science properties contribute novel causal powers (or at least novel groupings of causal powers), then such properties are natural\textsubscript{CS} but not natural\textsubscript{FC}.

Characterizing the debate about reduction as a contrast between views of natural properties also makes it clear why the debate between reductive and non-reductive physicalism is important. For one thing, we get a unified treatment of the questions about reduction that Bennett identifies as interesting. If non-reductive physicalism is true and mental properties must be included in the set of natural\textsubscript{CS} properties in addition to restricted-sense physical properties, then they “survive” multiple realization, and mental terms can refer to these properties and not their realizers. Reductive physicalism delivers the opposite verdict on these issues. Further, it is also clear why the debate has methodological implications. For instance, the reductive physicalist’s view on natural properties provides a particular interpretation of how to go about
locating mental properties in the natural world. The reductive physicalist must locate mental properties within the restricted-sense physical domain, on pain of eliminating them from causal and scientific discourse. If there are no irreducible natural mental properties, then this has important implications for the legitimate means of studying the mind – generalizations in terms of neurophysiological properties are arguably the highest level that would be legitimate (see note 11).

Finally, if more than the restricted-sense physical properties are natural_{CS}, then more abstract cognitive and social models of the mind can be developed relatively independently of neurobiology. Non-reductive physicalists might claim that this also fits nicely with the only evidence we have for physicalism – the actual success of the physical sciences in causally accounting for physical occurrences. The success we have had in causally explaining the vast majority of physical phenomena (like the development and movement of animals, the functions characteristic of life, and the products of chemical reactions) have not come from fundamental physics. In order for such successes to be evidence for physicalism we need to invoke a relatively broad notion of physical science, that is, any science that is compatible with \textit{fundamental category completeness}. In this sense, contemporary biology, chemistry, and psychology are all physical sciences (i.e., physical in the inclusive sense). Of course, reductive physicalists believe that they can account for this evidence equally well. This paper has not attempted to determine which view is true. But if the proposal put forth here is correct, it is clear why the debate is both philosophically and scientifically important.
References


