Supervenience-based Formulations of Physicalism*

JESSICA WILSON
University of Michigan

0 Introduction

The many and varied formulations of physicalism instantiate the following schema:

*Physicalism*: All entities are nothing over and above physical entities.

Filling in the schema requires specifying what it is for an entity to be physical, and what it is for an entity to be “nothing over and above” some other entities.¹ Some have worried that no account of the physical is adequate for physicalist purposes; and I’ll soon say a bit about how physicalists have responded (in my view, successfully) to this worry. But my main focus here is on nothing over and aboveness, and specifically on whether any supervenience-based approaches to characterizing this notion can enter into viably formulating physicalism.

Supervenience-based accounts of nothing over and aboveness also instantiate a schema:

*Supervenience-based Nothing Over and Aboveness*: The $A$-entities are nothing over and above the $B$-entities if the $A$-entities supervene on the $B$-entities.²

The four main ways of filling in the schema correspond to different ways of characterizing the modal strength, the supervenience base, and the supervenience connection at issue. For each such approach, I’ll argue that a physicalism based on the associated account of nothing over and aboveness is compatible with physicalism’s best traditional rival: a naturalist emergentism. Others have argued that supervenience-based formulations of

© 2005 Blackwell Publishing Inc., 350 Main Street, Malden, MA 02148, USA, and P.O. Box 1354, 9600 Garsington Road, Oxford, OX4 2DQ, UK.
physicalism fail. My aim here, besides addressing the full spectrum of supervenience-based approaches, is to show how certain philosophical and scientific theses concerning naturalism, properties, and laws give us new reasons to deny that supervenience is sufficient for nothing over and aboveness; and hence new reasons to think that supervenience-based formulations of physicalism are untenable.

1 Preliminaries

1.1 The physical/non-physical distinction

To set the stage for what follows, it’s worth seeing why the physical/non-physical distinction poses no deep problem for formulating physicalism. Physicalists generally cash out the distinction by reference to physics (by which is meant fundamental physics). For example, Hellman and Thompson (1975) say:

Mathematical physics, as the most basic and comprehensive of the sciences, occupies a special position with respect to the over-all scientific framework. In its loosest sense, physicalism is a recognition of this special position. (p. 551) [ . . . ] A thesis that qualifies as ontological physicalism [ . . . ] asserts, roughly, that everything is exhausted—in a sense to be explained—by mathematical-physical entities, where these are specified as anything satisfying any predicate in a list of basic positive physical predicates of [the relevant object language] L. Such a list might include, e.g., ‘is a neutrino’, ‘is an electromagnetic field’, ‘is a four-dimensional manifold’, and ‘are related by a force obeying the equations (Einstein’s, say) listed’, etc. (pp. 553–4)³

A physics-based approach to characterizing physical entities reflects, in part, traditional materialism’s evolution into physicalism. As Crane and Mellor (1990) tell the story, materialists specified characteristics definitive of matter—being impenetrable, being conserved, being such as (only) to deterministically interact, and so on; then claimed that everything was (nothing over and above the) material. But contemporary physics has shown that matter has few, if any, of these characteristics. Hence materialism has evolved into physicalism, marking a transition from an a priori to an a posteriori specification of the entities serving as a basis for ontological explanation. In particular, what entities are physical is to be determined by physical science (more precisely, physics) alone.

The main concern with using such an account to formulate physicalism goes under the rubric of “Hempel’s Dilemma” (acknowledging Hempel 1979).⁴ The first horn is straightforward: if the physics at issue in the physics-based account is current physics, then the resulting physicalism will surely be false; for current physics is surely both incomplete and to some extent inaccurate. There are three versions of the second horn, to the effect that if the physics at issue is rather future (or ideal) physics, then the
resulting physicalism will be either indeterminate in content (Hempel; Hellman 1985), trivially true (Chomsky 1968, Crook and Gillett 2001), or compatible with entities that are intuitively physically unacceptable (Papineau 1993, Loewer 2001). Granting that the difficulties with the first horn are intractable, only the third version of the second horn poses any real problem for formulating physicalism. An appeal to future physics does not render physicalism an indeterminate thesis, for the connotations of (fundamental) physics as a science treating of entities at relatively low orders of complexity will give the resulting physicalism at least some determinate content. The same connotations suffice to show that an appeal to future physics does not render physicalism trivially true, for not all entities will be treated by future physics (in particular, entities at high orders of complexity will not be); it is simply wrong to think of physics as the “science of everything”.

By way of contrast, the possibility that future physics might posit entities that are intuitively physically unacceptable—most crucially, that involve fundamental mentality—represents a genuine problem with using (future) physics alone to define the physical. If physics were to posit fields or particles themselves possessing mental properties (as opposed to composing or constituting complex entities possessing such properties), then physicalism would be—or should be—thereby falsified; for on any plausible historically grounded understanding, physicalism is incompatible with panpsychism (better: proto-psychism), the view that mentality exists at relatively low levels of constitutional complexity (that is, those levels treated by fundamental physics).

This difficulty may be avoided, however, by noting that Crane and Mellor’s genealogy omits a crucial fact: that physicalists have not handed over all authority to physics to determine, a posteriori, what is physical. Reflecting the historical roots of physicalism in materialism, as foundationally committed to understanding mentality as nothing over and above complex material goings on, one feature has remained definitive of the term “physical” (as this term enters into formulating physicalism, at any rate); namely, that physical entities are not fundamentally mental: physical entities do not individually either possess or bestow mentality. Hence a physics-based account of the physical should not be understood as the view that any and all entities treated by physics—current, future, or ideal—are physical. It should rather be understood as follows: an entity is physical just in case it is (approximately accurately) treated by current or future (at the end of inquiry, ideal) physics, and is not fundamentally mental. Positing the physicality of non-fundamentally-mental entities treated by better versions of physics prevents physics’ present failures from immediately falsifying physicalism, while providing continuous content to the account of the physical through the needed revisions. Most physicalists and non-physicalists have something like this account in mind (see Lewis’s
1983, p. 34 remark that the physical theory in physicalism is “something not too different from present-day physics, though presumably somewhat improved”) though they do not always make the relevant respect of similarity explicit. This response is, I believe, successful. Physicalists are not in trouble as regards the physical/non-physical distinction: neither the truth nor the question of physicalism turn upon it. (The afore-noted exception is that physicalism would be falsified if physics were to posit fundamentally mental entities.)

1.2 Same-subject necessitation

If there is a problem with formulating physicalism, it lies in the nothing/something over and above distinction. In investigating supervenience-based accounts of nothing over and aboveness, it will be convenient to focus on the central cases in the physicalism debates: those of same-subject necessitation, where two properties are instanced in a single subject, and one necessitates the other, with at least nomological necessity. A stock example is where an instance of a brain property in a person necessitates an instance of a mental property in that person; another is where an instance of a micro-structural property having such-and-such molecular lattice structure in a substance necessitates an instance of being fragile in that substance.7

Debates over these cases usually presuppose that the necessitating property P is nothing over and above physical properties (is ‘physically acceptable’, for short). Where P is physical, this presupposition is obviously satisfied; but more usually P will be of an entity not treated by physics (e.g., a brain), hence won’t be physical, properly speaking. In such cases, the presupposition requires an account of nothing over and aboveness on which, e.g., micro-structural properties are nothing over and above physical properties. Here I’ll grant that such an account is in place; if it is not, then physicalism faces difficulties in formulation besides those I canvas here. (That said, I’ll often use ‘physical’ to mean ‘physical or physically acceptable’.) Another presupposition is that, where necessitating property P is nothing over and above physical entities, it suffices for showing that necessitated property M is nothing over and above physical entities to show that M is nothing over and above P. This is plausible, even if (as will often be) two accounts of nothing over and aboveness are in play. Debate then proceeds over whether any same-subject necessitated properties are over and above their (physically acceptable) necessitating properties.

Cases of same-subject necessitation are central, first, because they are arguably the simplest cases where the question whether there are properties over and above physical properties comes up; as usual, it’s best to start simply, ignoring externalist and other complications. Second, the problem of higher-order (including mental) causation—of whether higher-order properties can be causally efficacious, given their strong dependence on lower-order (ultimately physical) properties—is most pressing for these
cases, for it is here that the higher- and lower-order properties (being seemingly distinct properties of the same subject, seemingly capable of producing the same effects) are in direct competition. The promise of resolving this problem is one of the primary motivations for physicalism. Cases of same-subject necessitation are thus a natural starting point for assessing the accounts of over and aboveness entering into a given formulation of physicalism; conversely, without an adequate account for these cases there might be no question of physicalism, after all.

1.3 Criteria of adequacy
I aim to determine whether a supervenience-based account of nothing over and aboveness can support a viable formulation of physicalism, applicable to the central cases. So I will assume that an adequate account of nothing over and aboveness must not render physicalism trivially true, trivially false, meaningless, or question-begging for these cases. I also impose two additional requirements. First, an adequate account must render physicalism a thesis that contrasts with its best traditional rivals—most importantly, a naturalist emergentism; I call this ‘the criterion of appropriate contrast’. Second, an adequate account must render physicalism a thesis that contrasts with (in particular) emergentism in an illuminating way—that is, in such a way as to avoid immediate stalemate (to ward off the debate’s being grounded in a clash of brute intuitions, or in a distinction that is irrelevant to the debate); I call this ‘the criterion of illuminating contrast’.

1.4 The coherence of emergentism
Emergentists are property (not substance) dualists, who maintain that certain structures composed wholly of physical entities have properties emergent from their necessitating physically acceptable properties. Commonly, emergent properties are understood as having or bestowing causal powers grounded in “configurational” forces or interactions that are as metaphysically and scientifically basic as the fundamental physical forces and interactions. When emergent properties are instanced in a particular, the operative forces are a combination of physical and configurational forces, and the particular thereby has causal powers going beyond those due just to the operative physical forces. These will generally include powers to move in ways affecting the physical entities composing the particular; hence emergentism is committed to the nomological possibility of “downward causation”.

Is emergentism, so conceived, coherent? If not, then no need for physicalism to contrast with it. Here I’ll present, in necessarily short order, responses to what I take to be the three main worries on this score. Along the way, I’ll point towards more detailed discussions of these responses, and provide further evidence of emergentism’s intended contrast with physicalism.
The first worry concerns whether emergentism is logically incompatible with contemporary physics. In response, McLaughlin (1992) has convincingly argued that emergent configurational forces are compatible with the laws and conservation principles of physics (see also Horgan 1993 and Papineau 2001). For example, Newton’s second law of motion, \( F = ma \), is neutral as regards what forces enter into the net force \( F \); hence it is compatible with these including a fundamental configurational force. Similarly for Schrödinger’s equation, \( \dot{\psi} = \frac{i}{\hbar} \mathcal{H} \), into which is inserted the Hamiltonian \( \mathcal{H} \) specifying the energies of the state (forces and energies being intertranslatable): “It is not that British Emergentism is logically incompatible with nonrelativistic quantum mechanics. It is not. Schrödinger’s equation could be the fundamental equation governing motion in a world with energies that are specific to types of structures” (p. 54). Configurational forces needn’t contravene the physical forces; rather, both kinds could operate in tandem, just as the diverse physical forces do. Nor are configurational forces incompatible with conservation laws, such as the relativistic principle of conservation of mass-energy: “[C]onfigurational forces need not involve any violation of this principle. [ . . . ] Configurational forces could involve various compensating shifts in mass and energy that maintained conformance to the principle of mass-energy” (p. 74).

In fact (as I discuss in Wilson 2002), recent scientific theorizing provides a blueprint for how conservation laws might support the warranted posit of emergent configurational forces. In the 1930’s, the law of conservation of mass-energy appeared to be violated in nuclear \(/^{12}_C\) decay interactions. Rather than accept the apparent violation as genuine, physicists posited a new fundamental force, or interaction—the weak nuclear interaction—as carrying away the missing energy. The phenomena to be accounted for in this case are not mental, so this new force, and its attendant properties and powers, were properly deemed physical. But were a similar apparent violation to occur in cases involving some mental phenomenon (and recalling that physical entities do not fundamentally have or bestow mentality), it is consistent with scientific practice that a new force be posited, of the sort that would falsify physicalism.

The second worry is directed at the emergentist’s claim that necessitated emergent properties have or bestow causal powers that their necessitating properties don’t have. It’s first useful to see how this claim is motivated by the problem of higher-order causation, as it arises in cases of same-subject necessitation. The immediate problem is a threat of causal overdetermination, where, implausibly, a single effect is caused twice over: once by the necessitating property \( P \), and again, independently, by the necessitated property \( M \). Physicalists and emergentists have distinct strategies—in my view, the most promising strategies—for avoiding this threat. (See Wilson 1999 for details.) Physicalists maintain that, for all \( M \) and \( P \) as above, every causal power of \( M \) is identical with a causal power of \( P \). This is compatible with \( M \)’s being distinct from \( P \) (if, say, \( M \) has a proper subset of \( P \)’s causal
powers), but in any case avoids overdetermination insofar as $M$’s causing of the effect is not distinct from $P$’s causing of the effect. Emergentists, in contrast, maintain that for some $M$ and $P$, $M$ has a causal power that $P$ doesn’t have. This avoids overdetermination in that $M$, not $P$, causes the effect.

While the emergentist and physicalist strategies for resolving the problem of higher-order causation seem both straightforward and straightforwardly distinct, one may worry that it is not possible for a same-subject necessitated property to have a causal power its necessitating property doesn’t have. As Kim (2000, p. 32) puts it: “The critical question [. . . ] is this: If an emergent, $M$, emerges from basal condition $P$, why can’t $P$ displace $M$ as a cause of any putative effect of $M$?”. Suppose, as is common enough, that a property’s having a causal power is a matter of nomological sufficiency (in the circumstances; henceforth assumed) for an effect of type $e$ (and where the circumstances alone are not nomologically sufficient for an effect of type $e$). Now consider the emergentist claim that $M$, an emergent property, has a causal power that $P$, its necessitating property, doesn’t have. Since $P$ necessitates $M$ with at least nomological necessity, $P$ is nomologically sufficient for $M$; and since $M$ is nomologically sufficient for an effect of type $e$, then (by transitivity of nomological sufficiency) so will $P$ be nomologically sufficient for an effect of type $e$. Hence $P$ will have any causal power that $M$ has, contrary to the emergentist claim.

One might deny that the having of a causal power involves nomological sufficiency or some other transitive relation (as does Armstrong 1983, p. 156), but even granting that it is plausible that $P$ inherits $M$’s causal powers, emergentists have (at least) two ways to respond. First, they can maintain that, while $P$ has every causal power $M$ has, these causal powers, as had by $P$, are not implemented or manifested in the same way as the causal powers, as had by $M$ (along lines, for example, of $P$’s being metaphysically, if not temporally, antecedent to $M$ in a causal chain leading to the effect). This strategy still avoids overdetermination (links in causal chains don’t overdetermine each other), and still contrasts with the physicalist strategy, in which $P$’s causing of the effect is identical with $M$’s.

Second, emergentists can respond by appealing to the thesis (which they traditionally recognize) that causal powers are grounded in specific fundamental forces or interactions. While this thesis is strangely ignored by contemporary philosophers, it is plausible, even obvious: the causal powers of being positively charged are grounded in the electromagnetic force; the causal powers of quantum color properties are grounded in the strong force; the causal power of being able to sit on a chair without falling through it is grounded (at least) in the gravitational and electromagnetic forces; and so on. Presumably, then, it makes sense to speak of the causal powers of a property relative to a particular set of forces. In particular, it makes sense to speak of the causal powers that a property has relative to the set of fundamental physical forces (these being, on the operative account of the physical,
the fundamental forces posited in physics, with the proviso that physical forces cannot involve fundamental mentality). Of course, physicalists think that physical forces are the only fundamental forces there are, while emergentists think that, in addition, there are one or more non-physical “configurational” fundamental forces. Emergentists can thus grant that, taking both physical and non-physical forces into account, $P$ has every causal power $M$ has; but coherently maintain that, when $M$ is emergent, it will have causal powers that are “new” relative to those powers of $P$ grounded only in fundamental physical forces. (See Wilson 2002 for further discussion.)

The third worry concerns emergentism’s commitment to the thesis of downward causation. Wouldn’t the efficacy of emergent properties vis-à-vis physical or physically acceptable effects violate the causal closure of the physical: the principle that every physical (physically acceptable) effect has a purely physical (physically acceptable) cause? And isn’t the causal closure of the physical widely accepted? Yes, and yes. As attention to the first worry illustrated, however, causal closure is not a principle of contemporary physics (though no doubt many scientists believe it). The acceptance of this principle is rather a constraint on physicalist theorizing (which motivates, in particular, the physicalist approach to the problem of higher-order causation); hence that emergentists deny it is no strike against their view.

I will assume, then, that a naturalist emergentism is coherent, scientifically respectable, and fundamentally at odds with physicalism. An account of nothing over and aboveness adequate to formulating physicalism must (in combination with the operative account of the physical) rule out the actual existence of emergent properties.

2 Supervenience-based accounts of nothing over and aboveness

Supervenience is a relation aimed at capturing the dependence of one family of properties on another by means of correlations alone, rather than by spelling out the precise nature of this dependence. Accounts of supervenience are most commonly directed at cases of same-subject necessitation. For example, “strong” supervenience is usually formulated as holding between families of properties $A$ and $B$, elements of which are co-instanced in individuals in a domain $D$:

$$A \text{ strongly supervenes on } B \text{ iff } \Box(\forall x \in D) (\forall a \in A) (x \text{ has } a \rightarrow (\exists b \in B) (x \text{ has } b \wedge \Box (\forall y \in D) (y \text{ has } b \rightarrow y \text{ has } a))).$$

“Weak” supervenience is defined similarly, modulo the absence of the second necessity operator. And “global” supervenience, on which $A$ globally supervenes on $B$ iff worlds with the same distribution of $A$-properties have the same distribution of $B$-properties, is also commonly directed at cases of same-subject necessitation: as Kim (1984, p. 68) says, “[W]ether two worlds
are discernible or indiscernible psychologically (or physically, etc.) is essentially a matter of how psychological properties are distributed over the individuals of the two worlds”.

That accounts of supervenience are standardly directed at cases of same-subject necessitation is unsurprising, given the centrality of these cases to the physicalism debates and the fact that, as Horgan (1993, p. 556) notes, “A rather dominant tendency since the early 1970’s has been to invoke [supervenience] in efforts to articulate a broadly materialistic, or physicalistic, position in philosophy of mind or in metaphysics generally”.

Davidson (1970, p. 88) is commonly cited as initiating this tendency in a well-known passage discussing his “anomalous monism” version of physicalism:

Although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events exactly alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect. Dependence or supervenience of this kind does not entail reducibility through law or definition.

Later Davidson (1973, pp. 716–7) explicitly advocated supervenience as a psychophysical determination relation: “Although, as I am urging, psychological characteristics cannot be reduced to the others [. . . ] there is a sense in which the physical characteristics of an event (or object or state) determine the psychological characteristics: in G. E. Moore’s word, psychological concepts are supervenient on physical concepts”.

As an editor pointed out, anomalous monism itself ultimately relies on a conception of nothing over and aboveness as identity (of events). Still, Davidson’s remarks were widely read as suggesting that psychophysical supervenience renders mental properties physically acceptable (see Horgan 1993, §IV for further discussion). The suggestion raised the hope, especially among philosophers concerned that reduction-based formulations of physicalism threatened the ontological and causal autonomy of higher-order properties, that supervenience correlations provide a way of ensuring that such properties are nothing over and above their physical base properties (in virtue of the correlations being sufficiently strong) while maintaining their ontological and causal autonomy (in virtue of these correlations being abstractly characterized, so as not to entail the reducibility of higher-order to base properties). This hope is broadly marked in our schema:

Supervenience-based Nothing Over and Aboveness: The \(A\)-properties are nothing over and above physical properties if the \(A\)-properties supervene on physical properties.
2.1 Supervenience simpliciter?

The general idea behind Davidson’s appeal to supervenience is that entities with the same physical properties must have the same non-physical properties.¹⁵ Let’s call supervenience so broadly characterized “supervenience simpliciter”, and fill in the schema:

The $A$-properties are nothing over and above physical properties if the $A$-properties supervene simpliciter on physical properties.

The suggestion that supervenience simpliciter can serve as a basis for formulating physicalism has been widely criticized. Schiffer (1987), Kim (1990), Heil (1992), Horgan (1993) and others have argued, largely on historical grounds, that the thesis that all properties supervene simpliciter on physical properties is consistent with non-physicalist views. Schiffer (p. 153), for example, notes: “[P]hilosophers who abhorred Moore’s irreducibly non-natural properties knew he also held this thesis about them: that it was not possible for two things or events to be alike in all physical respects while differing in some moral property. No one thought that Moore’s positive theory of moral properties was in any way mitigated by this further supervenience thesis”.

Philosophers (e.g., Witmer 2001, p. 63) sometimes dismiss this latter objection to using supervenience to formulate physicalism, on grounds that it is unclear that non-naturalism is coherent. The same cannot be said for emergentism, however; and emergentists arguably thought that emergent properties satisfied the correlations at issue in supervenience simpliciter. Broad (1925, pp. 67–8), for example (writing before the advent of quantum mechanical explanations of what he thought were emergent chemical properties), said:

No doubt the properties of silver-chloride are completely determined by those of silver and of chlorine; in the sense that whenever you have a whole composed of these two elements in certain proportions and relations you have something with the characteristic properties of silver-chloride, and that nothing has these properties except a whole composed in this way. But the law connecting the properties of silver-chloride with those of silver and of chlorine and with the structure of the compound is, so far as we know, an unique and ultimate law.¹⁶

Similarly for emergent properties of organisms:

[N]o amount of knowledge about how the constituents of a living body behave in isolation or in other and non-living wholes might suffice to enable us to predict the characteristic behavior of a living organism. This possibility is perfectly compatible with the view that the characteristic behavior of a living body is completely determined by the nature and arrangement of the chemical compounds which compose it, in the sense that any whole which is composed of
such compounds in such an arrangement will show vital behaviour and that nothing else will do so.\textsuperscript{17}

Though it is commonly assumed that the emergence of chemical or biological (non-mental) properties is no longer a live possibility (hence the contemporary focus, in the physicalism debates, on the ontological status of mentality), it’s clear from Broad’s remarks that nothing prevents an emergentist from maintaining that emergent properties supervene \textit{simpliciter} on physical properties. Hence a supervenience \textit{simpliciter} account of nothing over and aboveness violates the criterion of appropriate contrast.

2.2 \textit{Supervening with metaphysical necessity?}

The above considerations are suggestive but not decisive against supervenience-based formulations of physicalism, for historical appeals to supervenience are not specific regarding the modal strength of the correlations at issue; nor is supervenience \textit{simpliciter}. Though emergentists and physicalists agree that all properties supervene on physical properties, couldn’t their positions be distinguished by differing strengths of the supervenience correlations? After all, some emergentists (e.g., Mill) understood emergence to be a causal phenomenon; and it is commonly believed that causal relations hold with only nomological necessity (that is, hold only in possible worlds, so to speak, with the same laws of nature). Physicalists, by way of contrast, have usually taken the relations at issue in cases of same-subject necessitation to be ones—e.g., identity, parthood, or the determinable/determinate relation—that hold with metaphysical necessity (that is, hold in all possible worlds). So we might fill in the schema as follows:

The $A$-properties are nothing over and above physical properties if the $A$-properties supervene with metaphysical necessity on the physical properties.

Van Cleve (1990, p. 222) suggests something in this ballpark: “We can define emergence as follows: If $P$ is a property of $w$, then $P$ is emergent iff $P$ supervenes with nomological necessity, but not with logical necessity, on the properties of the parts of $w$”. Van Cleve’s remark that “the shape of an object is a logical consequence of the nature and arrangement of its parts” indicates that by “logical necessity” he means just metaphysical necessity, understood as above.

But while many philosophers think so,\textsuperscript{18} the contrast between metaphysical and nomological necessity doesn’t support an adequate account of nothing over and aboveness. For all properties could supervene with metaphysical necessity on physically acceptable properties, and yet some supervenient properties be physically unacceptable.
To see this, let’s first get clear on what metaphysical necessity comes to, and on how it differs from logical and nomological necessity. Witmer (2001, pp. 60–1.) provides a good starting point:

There are [...] various ways in which one might understand “logically possible.” On one reading, any sentence that fails to contradict itself given the rules of meaning governing its terms describes a logically possible situation. [...] A supervenience thesis of such strength would be far stronger than is necessary to express physicalism. Physicalism has long been understood to be a conceptually contingent claim [...]. There is another way of understanding the totality of possible worlds, however, that is of more relevance for physicalism. I have in mind that modality that is sometimes called “metaphysical possibility” or “broadly logical possibility”. The metaphysically possible worlds may be positively characterized as those which conform to facts about the natures or essences of entities in those worlds. That is, the metaphysically possible is what you get when you restrict the merely coherent by principles regarding what it is to be thus and such a thing. [...] There is nothing mysterious about this beyond what is mysterious about essences or nature—which is, it must be confessed, admitting a fair chunk of mystery. Claims about the “nature” of things seem part and parcel of the physicalist position, however, because it is, after all, a claim about the nature of the world.

These remarks indicate how to distinguish metaphysical from logical possibility. But where metaphysical possibility reflects the nature of things, what becomes of the distinction between metaphysical and nomological possibility in cases where the nature of a thing depends on the actual laws of nature? Obviously, it collapses: in such cases, what is metaphysically possible (or metaphysically necessary) for the thing just is what is nomologically possible (or nomologically necessary) for the thing. The present relevance of this is that there are good philosophical and scientific grounds for thinking that the natures of the entities under discussion in the physicalism debates indeed depend on the actual laws of nature.

The usual philosophical motivation for thinking this proceeds via the view that broadly scientific properties are essentially individuated by their causal features, including the causal powers they have or bestow. Then, insofar as laws of nature express (perhaps among other things) such causal potentialities, it follows that scientific properties are essentially individuated by their actual governing laws. This view is often called “necessitarianism about laws”; somewhat misleadingly, since it doesn’t entail that the actual laws hold in every possible world. “Necessitarians” generally allow that there could be worlds with completely different laws (so long as these contain only “alien” entities), and moreover allow for some contingency in what laws might govern actual scientific properties (e.g., that attending to constants dependent on contingent initial conditions, or more generally, that corresponding to laws very similar to, though not identical to, the
actual laws). For present purposes, the significant content of the view is as follows:

*The Necessitarian view:* Any possible world where there exists a scientific property of a type that actually exists is a world where hold all the laws actually governing that property (or laws very similar to these).

If this view is correct, then if there are emergent properties same-subject necessitated by physical properties in the actual world, it follows (since the latter properties are essentially individuated, in part, by reference to laws involving the former) that any world where the physical base properties are instanced will also be one where the emergent properties are instanced. Emergent properties will thus supervene with metaphysical necessity on their physical base properties, and a “supervening with metaphysical necessity” account of nothing over and aboveness will violate the criterion of appropriate contrast.

But many philosophers reject the Necessitarian view, instead maintaining that it is metaphysically possible that scientific properties be governed by laws very different from those actually governing them.\(^{20}\) This view is often called “contingency about laws”; which is again somewhat misleading, given previous remarks. The significant content of the view, for present purposes, is as follows:

*The Contingency view:* There are possible worlds where scientific properties of the type that actually exist are governed by very different laws than those actually governing such properties.\(^{21}\)

It’s worth observing, however, that many physicalists also consider themselves naturalists; and as I’ll now argue, the Contingency view is in considerable tension with naturalism. Naturalist philosophers have good *prima facie* reason to reject this view, and to rather maintain that the nature of scientific properties depends on the actual laws of nature (or laws very similar to these).

### 2.2.1 Naturalism and the Contingency view

Naturalism is the view that philosophical theorizing about the natural world should be consonant with scientific facts and theses, concerning both scientific subject matter and scientific methodology. As Post (1999) puts it:

Naturalism [is] the twofold view that (1) everything is composed of natural entities—those studied in the sciences (on some versions, the natural sciences)—whose properties determine all the properties of things, persons included, abstracta (abstract entities) like possibilia (possibilities) and mathematical
objects, if they exist, being constructed of such abstracta as the sciences allow; and (2) acceptable methods of justification and explanation are commensurable, in some sense, with those in science.

To be a naturalist philosopher is not to look to science for specific answers to philosophical questions. Scientific investigations are not, after all, conducted at the generally more abstract level of philosophical investigations. However, to be a naturalist is to require that one’s philosophical methodology and results be ‘commensurable’ (connoting: having a common measure) or ‘consonant’ (connoting: being in agreement or harmony) with scientific facts and theses, as these are presently reasonably taken to be.

Two points relevant to the upcoming discussion are in order. First, one needn’t self-identify as a naturalist in order to believe that, when it comes to philosophical investigations into natural phenomena, the sciences are the philosopher’s most important—in particular, best-confirmed and most comprehensive—source of data. One might, that is, be a naturalist just about natural phenomena, as an instance of a more general commitment to philosophical theses’ being appropriately sensitive to the relevant facts and theses about the phenomena under investigation. Since physicalism is a thesis about scientific entities, this thesis, and any upon which its formulation depends, are natural candidates for naturalist constraint.

Second, a naturalist approach is compatible with some revisionary understanding of the relevant scientific facts and theses. But as the terms ‘commensurability’ and ‘consonance’ imply, the working presupposition on this approach is in favor of explaining, as opposed to explaining away, these facts and theses. This is partly due to one primary motivation for naturalism; namely, the incredible success of the sciences. Any philosophical account maintaining that a primary tool of scientific methodology is unwarranted, for example, faces the burden of explaining how an unwarranted methodology has resulted in such success. But more generally, the presupposition reflects a common conception of philosophical methodology according to which, other things being equal, philosophical accounts of a given feature of reality that do not require extremely revisionary understandings of wide ranges of actual practice (that bear upon this feature, in particular) are to be preferred to those that do require this. The more revisionary the account, the heavier the burdens incurred when it comes to (a) motivating the revisions and (b) explaining away the facts and theses that the account deems misguided or incorrect. What I will mainly be arguing in what follows is that the Contingency view, and in particular the principles and theses standardly cited in support of the view, are incommensurable with certain pervasive scientific facts and theses; along the way it will, I think, be clear that proponents of this view have not come close to discharging the burdens incurred in endorsing such revisionary understandings.
Now, how are philosophers to determine whether a given metaphysical thesis is commensurable with the relevant scientific facts and theses? One might start by considering what philosophically informed scientists have had to say about the matter. The question is: does the nature of an actual scientific property depend on its actual governing laws? Here philosophers might take a leaf from the book of Bohm (1957, p. 14):

[C]ausal laws are not like externally imposed legal restrictions that, so to speak, merely limit the course of events to certain prescribed paths [. . . ] rather, they are inherent and essential aspects of these things [. . . ] Likewise, the general mathematical laws of motion satisfied by bodies moving through empty space (or under any other conditions) are essential properties of such bodies, without which they could not even be bodies as we have known them. Examples of this kind could be multiplied without limit. They all serve to show that the causal laws satisfied by a thing [. . . ] are inextricably bound up with the basic properties of the thing which helps to define what it is.

As with the Necessitarian view of laws, there is no need to be implausibly strict when interpreting such essentiality claims. Presumably Bohm could allow the same entity to be governed by similar, if not identical, laws: if the gravitational constant undergoes some minor shift, this needn’t make for a world of entirely different particulars; and similarly for the properties entering into causal laws. Here again the significant content of the view is that, within some reasonable degree of variation, the natures of scientific entities (in particular, properties) depend upon their actual governing laws.

Bohm’s view receives strong support from the following two facts, pertaining to both ordinary experience and scientific practice. First, the usual reason for positing properties, in either context, is to track similarities and differences in the causal actualities and potentialities of substantial particulars: in the usual case, for a substantial particular to have a given scientific property is just for the particular to be able to engage (in appropriate circumstances) in certain causal interactions rather than others. In other words, properties, at least of the sort at issue in the physicalism debates, appear to be defined by reference to (fairly) specific causal laws. Second and most crucially, in both ordinary and scientific contexts, we neither experience nor posit properties as persisting through changes in their governing laws (again, taking into account minor changes in constants, etc.). On the contrary: whenever a substantial particular S comes to have different causal powers at \( t' \) than it did at a previous time \( t \) (when in circumstances of the same type), we uniformly assume that S came to have one or more different properties at \( t' \) than it had at \( t \), not that the properties S had at \( t \) came to be governed by different laws at \( t' \). This fact is one of the things that any account of the nature of scientific properties should explain. The Necessitarian view, according to which this nature depends on the actual
governing laws (or laws very similar to these) straightforwardly explains this fact; and to this extent is commensurable with scientific practice and theorizing.

Contingency theorists may assert that their view also explains this fact, in that they suppose that the laws governing scientific properties remain the same within a world. One might wonder why Contingency theorists are entitled to suppose this, but let’s put that issue aside here. In any case, since their view is distinguished by the claim that actual scientific properties could be governed by very different laws at different possible worlds, it moreover needs to be established that this claim (hence the view) is commensurable with scientific practice and theorizing. So let us now turn to the principles and theses usually cited in support of the Contingency view.

First, let’s consider the principles usually cited in such support: Hume’s principle that there are no metaphysically necessary connections between distinct existences, and the modal ‘principle of recombination’ (discussed below). Lewis (1986a, p. 87), for example, first appeals to Hume’s principle as motivating the principle of recombination:

[W]e need a new way to say [. . . ] that there are possibilities enough, and no gaps in logical space. To which end, I suggest that we look to the Humean denial of necessary connections between distinct existences. To express the plenitude of possible worlds, I require a principle of recombination according to which patching together parts of different possible worlds yields another possible world. Roughly speaking, the principle is that anything can coexist with anything else, at least provided they occupy distinct spatiotemporal positions. Likewise, anything can fail to coexist with anything else.

He then uses the principle of recombination to motivate the Contingency view (p. 91):

Another use of my principle is to settle—or as opponents might say, to beg—the question of whether laws of nature are strictly necessary. They are not; or at least laws that constrain what can coexist in different positions are not. Episodes of bread-eating are possible because actual; as are episodes of starvation. Juxtapose duplicates of the two, on the grounds that anything can follow anything; here is a possible world to violate the law that bread nourishes. So likewise against the necessity of more serious candidates for fundamental laws of nature [. . . ].

Armstrong (1989, p. 20–21) similarly cites Hume’s principle as inspiring the principle of recombination he accepts, though (unlike Lewis) he thinks that the entities available for recombination must be actual; and he agrees (p. 90) that the principle of recombination entails the Contingency view: “[T]he Combinatorialist must accept the thesis that the laws of nature vary in different possible worlds. This corresponds to the intuitions of many, if not all, philosophers”.

Supervenience-based Formulations of Physicalism 441
Intuitions aside, should naturalists feel comfortable with such support for the Contingency view? Consider Hume’s principle. Hume’s reasons for endorsing this principle derived from his acceptance of a limited set of acceptable forms of reasoning, which did not include inference to the best explanation (so that one is barred from so inferring to the existence of metaphysically necessary connections), and where the allowable forms (so Hume argued) fail to warrant belief in such connections. Contemporary scientists do not accept Hume’s epistemological constraints, however; in particular, inference to the best explanation is an acceptable method of scientific justification and explanation, if any method is. (Indeed, contemporary advocates of Hume’s principle don’t accept Hume’s constraints, either, but seem mainly to accept his principle as an interesting constraint on their theorizing, as does Lewis in the introduction to his 1986b.) Moreover, (as in the case of Bohm), it appears that scientists do infer the existence of metaphysically necessary connections between distinct existences, in so many words. For example, contemporary expositions of particle physics and field theory are rife with talk of “essentially determined” force laws (Aitchison and Hey, 1989, p. 42) and “compulsory” existences (Ryder, 1996, p. 2–3). Why not take these claims at face value, as indicating that it is part of the nature of the scientific entities at issue to enter into such laws, or give rise to such existences? In any case, Hume’s principle is based in denying what is arguably the primary tool of scientific methodology; hence a naturalist has good prima facie reason to reject this principle, and any views whose support rests upon its acceptance.

Is the principle of recombination, that more directly motivates the Contingency view, commensurable with scientific facts and theses? Not so far as I can tell. It seems likely that scientists would respond with an incredulous stare to Lewis’s (1986a, p. 88) claim that “[I]f there could be a talking head contiguous to the rest of a living human body, but there couldn’t be a talking head separate from the rest of a human body, that [ . . . ] would be a failure of plentitude. (I mean that plentitude requires that there could be a separate thing exactly like a talking head contiguous to a human body)”; and similarly for his claim that plentitude requires (with few exceptions) that any fundamental physical property might or might not be paired, as a matter of law, with any other. Of course, philosophers often say things that surprise scientists. The present point is simply that insofar as it is evident that scientists don’t modally reason in accord with the principle of recombination, it remains deeply unclear how this principle is to be reconciled with scientific practice. Hence a naturalist also has prima facie reason to reject this principle, and the Contingency view it entails.
Second, consider Armstrong’s (1983, p. 158) motivation for the Contingency view, which appeals to the range of scientific hypothesizing:

[1] In trying to discover the laws of nature, scientists feel free to consider possibilities in a very wide-ranging manner, quite unlike the constraints which naturally suggest themselves in logical and mathematical argument. It would have to be admitted, at the least, that the laws of nature give a definite impression of contingency.

It is certainly true that scientists consider a range of hypotheses when attempting to discover what laws govern a given property, and a naturalist will appreciate Armstrong’s attempt to locate support for the Contingency view within scientific practice. But the attempt fails. First, it is unclear why these hypotheses should be understood as tracking the metaphysical possibilities for a given property, as opposed to merely epistemic possibilities, reflecting the investigator’s state of ignorance. Second, in any case the hypotheses scientists entertain concerning a property’s governing laws are not so “wide-ranging” as to support the Contingency view over the Necessitarian view. Presumably the hypotheses scientists consider will have to correctly track what happens when the property is instanced in various circumstances; but this will greatly constrain the range of available hypotheses, and hence the range of laws that Armstrong supposes will be indicated by the hypotheses. Indeed, correctly tracking the facts will provide at least one dimension along which the laws indicated by the hypotheses and the actual laws will be very similar; and since so-called Necessitarians can admit that properties of the type that actually exist might be governed by laws that are very similar to the actual laws, they can accept Armstrong’s observation without accepting his view. Stronger support for the Contingency view is needed, of the sort that Hume’s principle or the principle of recombination would give, if only they didn’t violate naturalist tenets.

Third and finally, the Contingency theorist might attempt to motivate their view via the claim that scientific properties have either an intrinsic aspect (e.g., a non-relational qualitative property, or a property of “this-ness”) or an intrinsic identity. For convenience I’ll focus on the latter suggestion; it should be clear that my remarks also apply to the former.

By analogy to *haecceities*—intrinsic identities of substantial particulars, such that substantial particulars may be the same in spite of having completely different properties—the Contingency theorist might suggest that properties have *quiddities*—intrinsic identities such that properties may be the same in spite of being governed by completely different laws. And indeed, both Armstrong and Lewis accept the existence of quiddities, so understood, as required to make sense of the Contingency view: Armstrong (1989, p. 44) says: “[P]roperties and relations do each have their own haecceity, or, better, their own quiddity or nature”, and Lewis stated his acceptance of quiddities in his 2001. But what grounds are there for thinking
that scientific properties have an intrinsic (Armstrong also says: primitive) identity, independent of their actual governing laws?

Armstrong (1983, p. 160) offers the following motivation:

[S]woyer [. . .] argues that properties must have ‘essential features’ [namely] the relations of ‘nomic implication’ which properties have to other properties. But why need properties have essential features at all? Perhaps their identity is primitive. To uphold this view is to reject the Principle of the Identity of Indiscernibles with respect to Properties. Properties can be different, in the same way that, many of us would maintain, ordinary particulars can just be different although having all their features in common [. . .] properties can be their own essence.

More to the present point, to allow that properties have an intrinsic identity (primitive or not) is to reject the Distinctness of Discernibles: properties can be the same, in spite of being governed by completely different laws, just as (those accepting haecceities—primitive “essences” of particulars—might maintain) ordinary particulars can be the same, in spite of having none of their features in common.

Armstrong’s argument for properties’ having intrinsic identities turns on a supposed analogy to substantial particulars’ having intrinsic identities; but the analogy, and so the argument, fails. First note that there is a case, in line with naturalist tenets, for thinking that some substantial particulars, at least, have intrinsic identities. This is because (as previously noted) inference to the best explanation is an accepted mode of inference in the sciences, and in the case of substantial particulars, there is something to explain—namely, our common experience of substantial particulars persisting through relatively extreme changes in their properties (as when a single human moves from infancy to adulthood)—for which the thesis that substantial particulars have intrinsic identities is the best, or at any rate a reasonable, explanation.

But—and here we return to the aforementioned crucial fact about properties—we do not, in either ordinary or scientific contexts, experience or posit properties as persisting through any but very minor changes in their governing laws (e.g., those attending a minor shift in a constant, say). So there is no motivation here, as there arguably is in the case of substantial particulars and their properties, for thinking that properties have an identity completely independent of their governing laws. There is nothing to explain, such that the thesis that properties have intrinsic identities would be the best, or at any rate a reasonable, explanation of it. So the analogy fails, and the Contingency view remains unsupported, by naturalist lights.

Perhaps there is some other motivation for the Contingency view, besides those considered here, that is naturalistically acceptable. What I hope to have shown is that, at least as usually motivated, the Contingency view is in
considerable tension with naturalism. By way of contrast, something in the proximate vicinity of the Necessitarian view appears perfectly consonant with the relevant scientific facts and theses. Naturalist physicalists thus have good prima facie reason to reject the Contingency view, and accept that the nature of scientific properties depends on their actual governing laws, or laws very similar to them, in which case a “supervening with metaphysical necessity” account of nothing over and aboveness will, as previously discussed, render the associated formulation of physicalism compatible with emergentism, and hence violate the criterion of appropriate contrast.

One might wonder: even supposing one rejects the Contingency view, might there be some principled way of maintaining that physical properties essentially depend on some but not all of their actual governing laws, and not, in particular, on any actual laws connecting physical properties with emergent properties? And if so, couldn’t the present supervenience-based account of nothing over and aboveness be salvaged, by requiring that the base properties be instanced in worlds with only physical laws? This strategy is more properly situated at the end of section §2.3; but a naturalist assessment of the strategy can start now.

First, the scientific facts and theses we have considered in this section don’t provide support for the claim, upon which the strategy depends, that any scientific properties are individuated by only certain of their actual governing laws. Once again, consider the crucial fact that we do not experience or posit properties as persisting through any but minor changes in their causal powers. This practice provides no support for thinking that any scientific properties are individuated by some, but not all, of their actual governing laws, and rather provides support to the contrary. And insofar as hypotheses concerning the laws that may actually govern a property need, at a minimum, to correctly track the facts about instances of the property, what ground is there for thinking that, at least sometimes, only some, but not all, of these facts need to be tracked, in determining a property’s governing laws? Relatively, why think that the laws governing any actual emergent properties there may be can be “hived off” of the laws governing their physical necessitating properties, while leaving the physical laws intact?

Second, if the remarks in the next section are correct, there are positive scientific (hence, naturalist) grounds for thinking that some actual laws can’t be hived off from other actual laws, in the way the strategy requires.

2.2.2 Holism about Natural Laws
Scientific grounds for thinking that emergent properties might supervene with metaphysical necessity on their necessitating properties stem from the scientific hypothesis that there are holistic constraints on the actual laws of nature, such that any world in which some of the actual laws hold is a world where all of them hold. Let’s call this hypothesis
Holism about Natural Laws: The actual laws of nature are such that any possible world in which some of the actual laws hold is a world in which all of them hold.

If this hypothesis is true, then again, emergent properties will supervene with metaphysical necessity on their base properties. I believe the hypothesis is most likely true; but it is enough for present purposes to show that the hypothesis is compatible with what scientists presently believe. And for this it suffices to consider one working assumption, presently directing thinking and theorizing about the nature of the laws governing fundamental interactions, that reflects an important sense in which scientists think the actual laws of nature are holistic.

The assumption I have in mind is arguably the driving hypothesis of contemporary fundamental physics, and should be boringly familiar: it is that the fundamental interactions are unified, in being distinct aspects of a single interaction. The current thinking is that at extremely high energies (around $10^{19}$ GeV, such as were present $10^{-41}$ seconds after the Big Bang), there was only a single force/interaction; distinct manifestations of this single interaction appear as energies decrease and various symmetries are broken. This hypothesis is motivated by, among other considerations, the fact that the coupling strength of the various interactions varies with the energies involved in the interaction, in such a way that theory-driven extrapolations from known variations indicate that the coupling strengths coincide at sufficiently high energies.

Here the language of “unity” makes it reasonable to take the hypothesized connection between the fundamental interactions, and hence between their associated laws, to hold with metaphysical necessity. Even so, this assumption is, in the nature of the case, directed at laws concerning the fundamental physical interactions (the electromagnetic, strong nuclear, electro-weak, and gravitational interactions). What reason is there to think the assumption would apply to laws concerning fundamental interactions giving rise to emergent properties?

My reasons are again founded in my thinking that the warranted posit of fundamental emergent laws/interactions would proceed in exactly the same way as the warranted posit of fundamental physical laws/interactions. (Recall the §1.4 analogy to the warranted posit of the weak interaction in the face of apparent violations of conservation laws during $\beta$ decay.) Since we can understand emergent laws and interactions along the same lines as physical laws and interactions, it seems reasonable to think that the assumption that the fundamental interactions are unified would extend to any emergent interactions there might be. If so, then scientific commitment to the unification of all interactions supports Holism, and undermines hopes that the distinction between nomological and metaphysical necessity, as a way of characterizing nothing over and aboveness, suffices to distinguishes physicalism from emergentism.
But what if physicists are wrong—the fundamental interactions aren’t unified, and Holism isn’t true? Then physicalism should still not be formulated in terms presupposing the falsity of this hypothesis, for the compatibility of Holism with what scientists presently believe is enough to show that the distinction between metaphysical and nomological necessity is irrelevant to the debate between physicalists and their rivals. Surely the viability of physicalism should not turn on whether or not the fundamental interactions are unified! Similarly, we should not rest physicalism’s viability on controversial or implausible principles, such as Hume’s principle or the principle of recombination. In either case, the resulting physicalism avoids violating the criterion of appropriate contrast only by violating the criterion of illuminating contrast.

2.3 Supervening on a restricted supervenience base?
A plausible diagnosis for the failure of the previous supervenience-based approaches to nothing over and aboveness is that these do not distinguish two importantly different ways in which properties may satisfy supervenience correlations. First, the correlations might be satisfied by instances of the base property’s completely determining instances of the supervenient property. Second, the correlations might be satisfied by instances of the base property’s being preconditions for instances of the supervenient property, which preconditions are necessarily correlated with further conditions needed for instancing of the supervenient property. Emergentism and occasionalism are each compatible with physical properties’ being preconditions necessarily correlated (with whatever modal strength one likes) with further conditions that necessitate the supervenient property in a physically unacceptable fashion. Hence it’s no surprise that satisfaction of supervenience correlations alone doesn’t guarantee that supervenient properties are physically acceptable.29

Several supervenience-based accounts of nothing over and aboveness attempt to guarantee the requisite complete determination, by imposing additional constraints on the characterization of the supervenience base.30 The constrained base is generally characterized by reference to “non-buttery” physical possibility. One proceeds first by specifying the physically possible worlds, usually by reference to physical laws of nature: a physically possible world is a world where hold all the physical laws that actually hold. One then attempts to cull down the physically possible worlds so as to rule out those which are also “buttery” in containing physically unacceptable entities, such as emergent properties or supernatural particulars. Nothing over and aboveness is then a matter of supervening on physically acceptable properties in all non-buttery, physically possible worlds. As a partial filling in of the schema:

The \( A \)-properties are nothing over and above physical properties if the \( A \)-properties supervene on physical properties in all non-buttery, physically possible worlds.
There are several accounts of what physically possible worlds are to be non-buttery. On Lewis’s (1983) account, these are physically possible worlds with no “alien” properties, where a property is alien to a world \( w \) just in case it is neither instanced in \( w \) nor analyzable in terms of properties instanced in \( w \). Lewis (p. 37) says: “If our world is Materialistic, then it is safe to say that some of the natural properties instantiated in any nonmaterialistic world are properties alien to our world”. He then formulates physicalism as the thesis that all properties supervene on physical properties in worlds relevantly similar to ours, and in which there are no properties alien to our world: “Within the inner sphere of possibility, from which these alien intrusions are absent, there is indeed no difference of worlds without a difference in their arrangements of qualities” (1986, x). On Lewis’s view, these qualities are determined by physics, so let’s go back to calling these qualities ‘physical’ or ‘physically acceptable’ properties. Filling in our schema:

The \( A \)-properties are nothing over and above physical properties if the \( A \)-properties supervene on physical properties in all physically possible worlds in which there are no alien properties.

Lewis’s account of nothing over and aboveness doesn’t work, however, for emergentists don’t think that emergent properties are alien. Moreover, emergentists agree that emergent properties supervene on physical properties, and (being good naturalists) they think that emergent, physical, and physically acceptable properties are the only sorts of properties there are. So emergentists will agree that all properties supervene on physical properties in nearby worlds in which there are no alien properties. Lewis’s specification of the relevant supervenience base doesn’t guarantee physicalism’s incompatibility with emergentism, and so violates the criterion of appropriate contrast.

But perhaps (revisiting his remark “If our world is Materialistic . . . ” above), Lewis intended his formulation of physicalism to be partially based on the hypothesis that the actual world is “Materialistic”, to the effect that (a) physicalism is actually true and (b) all properties supervene on physical properties in nearby possible worlds with no alien properties. But since emergentists and physicalists agree as regards (b), their dispute will lie in (a): emergentists will deny that physicalism is true in the actual world, and physicalists will assert this. Such a formulation of physicalism leads immediately to stalemate, and so violates the criterion of illuminating contrast.

On Haugeland’s (1983) account, non-buttery physically possible worlds are physically possible worlds where no physical laws are violated. Filling in:

The \( A \)-properties are nothing over and above physical properties if the \( A \)-properties supervene on physical properties in all physically possible worlds in which no physical laws are violated.
But emergent properties need not violate physical laws. As Horgan (1993, p. 557) remarks: “[Are] the laws of physics abrogated when emergent properties are instantiated? According to the emergentists, no. For the laws of physics do not actually assert that physical forces are always the only operative forces in a physical system. So the laws of physics remain true when an emergent property is instantiated”. As per the discussion in §1.4, emergentists are right about this. To repeat: the laws of physics are compatible with the existence of configurational forces or energies giving rise to emergent properties, and any such forces or energies could operate in tandem with the physical forces and energies. So Haugeland’s restriction doesn’t rule out worlds with physical laws like ours and, in addition, emergent properties.

Haugeland’s restriction also doesn’t rule out worlds with physical laws just like ours and, in addition, disembodied angels (or occasioning deities). He acknowledges this, and goes on to rule out angel worlds via the thesis that the mental supervenes on the physical: “[This thesis] amounts to a further culling of the set; it says that some of these worlds are not “really” possible after all, because they still do not bear a close enough relation to the physical” (p. 99). But such a thesis won’t rule out emergent worlds, since emergent properties do supervene on the physical.

On Hellman and Thompson’s (1975) account, non-buttery physically possible worlds are worlds with only physical laws. Filling in:

The A-properties are nothing over and above physical properties if the A-properties supervene on physical properties in all physically possible worlds with only physical laws.

The intuition that over and above properties invoke new systems of laws is a good one; for example, Broad characterized emergence in terms of what he called “trans-physical” laws. But Broad never assumed that worlds with only physical laws were possible. Rather, he used an epistemological criterion to get at the metaphysical distinction between laws, which criterion was applicable to laws all present at a world: trans-physical (emergent) laws governed phenomena unpredictable, even in principle, from the physical (physically acceptable) laws. Hellman and Thompson’s account, however, requires the possibility of worlds with only physical laws.

We are now back to worries regarding whether worlds with only physical laws really are possible. Again, by the lights of naturalist philosophy and contemporary science, the laws of nature—whether or not these include laws governing emergent properties—are arguably holistic. If so, and if emergentism is true, then there are no metaphysically possible (hence no physically possible) worlds with “only” physical laws. Hellman and Thompson’s account, like previous accounts, fails to guarantee that physicalism is incompatible with emergentism.
Similarly inadequate is Jackson’s oft-cited 1998 account. Here the physically possible worlds in the supervenience base are ‘minimal physical duplicates’ of the actual world (so that physicalism is the thesis that every minimal physical duplicate of the actual world is a duplicate *simpliciter*). Then: a physical duplicate duplicates the physical facts as well as the laws of physics, and a minimal such duplicate duplicates just this and *nothing more than is metaphysically necessary*.

Jackson does (p. 15–17) consider the objection that “physical properties are necessarily connected to non-physical properties, and so any minimal physical duplicate of our world is bound to have some non-physical nature”. But (as the “bound to” expression indicates), he has in mind seemingly non-physical properties that are logically entailed by physical properties: “For instance, having mass is a physical property [. . .] but it is necessarily connected to having mass or being made of ectoplasm, and having mass or being made of ectoplasm is not a physical property”. Jackson then makes a tripartite distinction between properties: there are the physical (or physically acceptable) properties whose actual existence is accepted by all participants to the physicalism debate, the non-physical properties whose actual existence would falsify physicalism, and what he calls “onlooker” properties, that do not seem to fall into either of the previous divisions, but whose existence is in any case acceptable to both physicalists and non-physicalists. After noting that *having mass or being made of ectoplasm* is an onlooker property (since it can be possessed either by having mass, which is fine by physicalists, or by being made of ectoplasm, which is fine by some non-physicalists—though not emergentists), Jackson says

> We can now give the reply to the objection from necessary connections between properties. None of the plausible examples of necessary connections from physical properties to distinct properties that are not physical properties is an example of a connection from a physical to a non-physical property. They are all like our example of having mass and having mass or ectoplasm: the necessary connections are between physical properties and onlooker properties.

Jackson provides no argument for the claims entering into this reply; and if I am right no such argument will be forthcoming. For above I provided plausible examples of necessary connections from physical to distinct non-physical (non-onlooker) properties. The examples I discuss are not the result of logical gerrymandering; rather, they involve metaphysically necessary connections, that hold in virtue of the natures of (a) scientific properties, or (b) the laws of nature. Once again for good measure: if physical properties are essentially individuated by the laws that actually govern them (or laws very similar to these), as is reasonably motivated by naturalist considerations, and if emergentism is true, then (case a) there will be metaphysically necessary connections between physical properties and
emergent properties. This will also be true (case b) if emergentism is true, and if the laws of nature are holistic, as is reasonably motivated by the scientific hypothesis that the fundamental interactions are unified.

Since such plausible examples are now on the table, Jackson and other supervenience physicalists cannot just insist that any properties supervening with metaphysical necessity on physical properties must be either physical (physically acceptable) or “onlooker” properties, without begging the question of whether supervenience is sufficient for physicalistic acceptability (and Jackson’s arguments, in particular, go nowhere towards establishing that he has a right to so insist). Rather, Jackson and others must address, and somehow reject, the specific naturalist and scientific reasons for thinking that there can be metaphysically necessary connections between physical and emergent properties. But how are naturalist physicalists to reject these reasons—or better yet, why? In any case, we can move on: appeals to restricted supervenience bases are clearly no improvement on previous supervenience-based accounts of over and aboveness.

2.4 Supervening via a strengthened supervenience connection?
What now? The last hope for a supervenience-based formulation of physicalism is to strengthen the supervenience connection. In Horgan’s terms, what a supervenience-based physicalism needs is “superdupervenience”: supervenience guaranteeing that supervening properties are nothing over and above their physically acceptable base properties. Horgan (1993, p. 563) suggested the following constraint for the job:

**Horgan’s Constraint**: Any genuinely physicalist metaphysics should countenance ontological inter-level supervenience connections only if they are robustly explainable in a physicalistically acceptable way.

Building the the requirement of “physicalistic acceptability” into the operative notion of robust explainability, our schema becomes:

The $A$-properties are nothing over and above physical properties if the $A$-properties supervene on physical properties, in such a way that the former are robustly explainable in terms of the latter.

As should be clear by now, Horgan is right that something beyond supervenience is needed to guarantee nothing over and aboveness for cases of same-subject necessitation. I have argued (Wilson 1999) that Horgan’s Constraint isn’t what’s needed, but here let me raise one different point, and one point reflecting Horgan’s negative assessment of the prospects for a superdupervenience-based physicalism.

First, accepting such a strengthening of the supervenience connection means giving up the reasons for characterizing nothing over and aboveness
in terms of supervenience in the first place. The hope, recall, was that a supervenience-based account of nothing over and aboveness offered an alternative to reduction-based accounts (which accounts were supposed to threaten the ontological and causal autonomy of same-subject necessitated properties). But what does “robust explanation” come to, if not functional or causal reduction?34 Horgan’s illustration (p. 579) of a supervenient property susceptible to a successful robust explanation supports such a reading:

Explaining why liquidity supervenes on certain microphysical properties is essentially a matter of explaining why any quantity of stuff with these microphysical properties will exhibit these macro-features [tendency to flow, to assume shape of vessel that contains it, etc.] [. . .] this suffices to explain the supervenience of liquidity because those macro-features are definitive of liquidity [and because] it seems explanatorily kosher to assume a “connecting principle” linking the macro-features to liquidity, precisely because those features are definitive; the connecting principle expresses a fact about what liquidity is.

Similar remarks apply to Chalmers’s (1996) formulation of physicalism in terms of “broadly logical” supervenience, which requires that a relation of entailment hold between the primary intensions of the concepts associated with the supervenient and base properties (for short, let’s say that the latter properties conceptually entail the former). The point of requiring such entailments is, of course, precisely that they guarantee that a functional or causal reduction of the supervenient to subvenient properties is available. For example, Chalmers says of heat: “The concept of heat that we had a priori—before the phenomenon was explained—was roughly that of ‘the thing that plays this causal role in the actual world.’ Once we discover [a posteriori] how that causal role is played, we have an explanation of the phenomenon” (45). Filling in:

The A-properties are nothing over and above physical properties if the A-properties supervene on physical properties, in such a way that the former are conceptually entailed by the latter.

Given the import of the robust explanations and conceptual entailments at issue, we can characterize Horgan’s and Chalmers’s suggestions as:

The A-properties are nothing over and above physical properties if the A-properties supervene on physical properties, in such a way that the former are functionally or causally reducible to the latter.

We have come full-circle through the spectrum of supervenience-based accounts of nothing over and aboveness, to return to a reduction-based
account. To be sure, functional or causal reductions differ in interesting ways from the sort of reductions, requiring type-identity or extensional equivalence of reduced and reducing entities, that physicalists originally appealed to supervenience to avoid. But functional or causal reductions also invoke worries about ontological and causal autonomy of functionally or causally reduced entities: after all, such reductions aim to establish that the “causal work” associated with the higher-order entity is being done by the lower-order property; and if the higher-order property is doing no independent causal work, why think it exists? Whether or not these threats can be avoided, in any case supervenience is not doing any independent work here. Hence even supposing a superdupervenience-based formulation of physicalism were satisfactory, it would be disingenuous to claim that this vindicated supervenience-based accounts of nothing over and aboveness.\(^3\)

Second, in any case there is reason to think that a superdupervenience-based formulation of physicalism is unsatisfactory. Horgan suspects, as does Chalmers and do I, that the explanatory gaps between qualitative mental and physical properties are intractable. A physicalist would be unwise, then, to formulate physicalism in terms of superdupervenience, for the absence of robust explanations or conceptual entailments for cases of qualitative mental properties appear to render such a formulation immediately, if not quite trivially, false.

### 3 A closing diagnosis

Having run through the spectrum of supervenience-based approaches to characterizing nothing over and aboveness, it seems safe to say that there is no question of supervenience-based formulations of physicalism. Why, then, have so many participants to the physicalism debates endorsed such formulations, especially (putting aside the non-starters discussed in §2.1 and §2.4) as involving supervening with metaphysical necessity (§2.2), perhaps first restricting the supervenience base (§2.3)?

Plausibly, proponents of these approaches have thought it relatively unproblematic to assume that the connection between physical properties and whatever emergent properties they might necessitate is modally weaker than that holding between physical and physically acceptable properties, such that it makes sense to assume that emergent properties are only nomologically necessitated by physical properties, and to assume, for example, that one can characterize the physicalist’s supervenience base by reference to worlds with only physical goings-on. I say “relatively” unproblematic because these proponents seem often to be aware that causal essentialist accounts of the nature of scientific properties (e.g., Shoemaker’s) and the concomitant Necessitarian view of laws are incompatible with these assumptions. But (and herein lies my diagnosis) since such accounts of properties and laws are controversial, those formulating
physicalism as a supervenience thesis have felt within their rights to set these glitches aside.

What I have attempted to show here is that this manoeuver is unjustified, and most importantly is in deep tension with the naturalistic motivations for both physicalism and its best rival, emergentism. That broadly scientific properties might be essentially individuated by all their actual governing laws (or laws very similar to these) is not just an outside philosophical position, but rather is strongly motivated by scientific practice (not to mention ordinary experience). Conversely, the view that scientific properties might be divorced from their governing laws is motivated by principles (such as Hume’s principle and the modal principle of recombination) that are either so epistemologically constrained that no scientist could afford to implement them, or so implausible that no scientist could stand to endorse them. Moreover, independently of concerns about the nature of properties, contemporary scientific theorizing supports taking seriously, at the least, the view that the laws of nature—whether or not these involve emergent properties—are holistic, such that blithe assumptions about the possibility of worlds with only physical laws are naturalistically suspect. When the topic of discussion is the ontological status of broadly scientific properties, such considerations cannot be set aside. If those endorsing supervenience-based formulations of physicalism are to be true to their naturalistic roots, they must directly grapple with the seeming incommensurability of their modal assumptions about properties and laws with scientific practice and theorizing.

Notes

* Many thanks to Richard Boyd, Janice Dowell, Gene Mills, Sydney Shoemaker, Jason Stanley, two referees for and the editors of *Nouš*, and (in greatest measure) Benj Hellie, for helpful discussion and/or comments on previous versions of this paper.

1 It also requires specifying the entities in the domain of the quantifier. Physicalism is usually understood as a comprehensive thesis about the actual world, that is contingently true, if true; so the domain presumably does not include all possibilia. Discussions of physicalism tend to focus on broadly scientific entities, initially ignoring, e.g., mathematical and metaphysical entities; I'll assume the restriction to broadly scientific entities here.

2 The schema contains an ‘if’ rather than an ‘iff’ clause to allow for the possibility that different accounts of nothing over and aboveness might be required to handle different kinds of relations between broadly scientific entities. The central cases that will be the focus of discussion (see §1.2) are those to which supervenience-based accounts should apply, if they apply to any (see §2).


4 The remarks to follow are necessarily abbreviated; a more thorough discussion of Hempel’s Dilemma and defense of a physics-based account of the physical against this and other concerns may be found in Wilson (forthcoming).

5 Nor should remarks (like those of Hellman and Thompson, above) that physics is "comprehensive" be interpreted as suggesting that physics is the science of everything. To say that physics is comprehensive is to say (as physicalists do) that entities that are not, properly
speaking, physical, are nonetheless nothing over and above ("exhausted [by]—in a sense to be explained") physical entities.

See Papineau 1993, Ravenscroft 1997, Papineau 2001, and Loewer 2001 for variations on this theme. See also Crook and Gillett 2001 for a relevantly similar alternative to physics-based accounts, on which the basic physical entities are, roughly, the contingent non-mental ontologically basic entities. One might wonder (as Janice Dowell did) how these accounts make sense of the type-identity physicalist’s claim that mental entities are identical with physical entities, which claim seems to entail (by Leibniz’s law) that some physical entities are both fundamental and mental. Sense is made by observing that the identity theorist’s claim is cast in loose language: strictly speaking, mental entities are identical to physically acceptable (not physical) entities (see §1.2); and physically acceptable entities (e.g., micro-structural properties) are not relatively fundamental, in the sense of ‘relatively ontologically basic’ at issue here.

I focus on (repeatable) properties for continuity with upcoming accounts, but the forthcoming remarks appear to go through, mutatis mutandis, if cases of same-subject necessitation instead involve events, facts, tropes, or truths; and also if they really involve multiple subjects (e.g., bodies and persons) or probabilification.

Protopsychism and substance dualism are also worthy traditional rivals; but as noted, physicalism’s contrast with protopsychism doesn’t turn on the nothing/something over and above distinction, and substance dualism isn’t obviously relevant to cases of same-subject necessitation between properties. Another worthy rival is Malebranchean occasionalism, according to which a supernatural deity brings about the instancing of higher-level properties upon the instancing of lower-order physically acceptable properties. Though supervenience also fails to distinguish occasionalism from physicalism, I won’t do much to establish this beyond the occasional aside. Another traditional rival is non-naturalism (as per Moore 1903), but since its coherence is controversial, I won’t discuss it here.

Prototypically, the British emergentists Mill (1843), Bain (1870), Morgan (1923), and Broad (1925).

Kane (1993, pp. 2–3) confirms: "\( F = ma \) is used to compute the motion of an object, given any force \( F \) on the object. And specific classical forces have been discovered, such as gravity with \( F = \frac{GmM}{r^2} \). Hamilton’s or Lagrange’s equations are equivalent to \( F = ma \) in a different formulation. In quantum theory there is an analogous structure. The Schrödinger equation \[ \ldots \] is like \( F = ma \). It holds for any Hamiltonian. Specific forces lead to specific Hamiltonians".

It’s worth noting that nuclei are composite entities, and thus that scientists had no problem with positing the existence of a fundamental “configurational” force. In fact the “nuclear” interactions are now understood as occurring between sub-nuclear entities; but the point remains: there is nothing incoherent about fundamental features arising from complex configurations.

Kim situates his worry here and elsewhere in a context where the putative effect \( M^* \) itself has a base property \( P^* \) (in which case, he argues, \( M \) must be downwardly causally efficacious vis-à-vis \( P^* \) if \( M \) is to cause \( M^* \)), but this is inessential. See Wilson 2002 for a more detailed presentation of the worry than I give here, in a general context.

It may be that the second strategy is best seen as offering a specific way of making sense of the first. See Wilson (in progress) for discussion.

Alternatively, global supervenience is used in order to relax the restriction to same-subject necessitation (see Melnyk 1991), as is the “regional” supervenience formulated and discussed in Horgan 1982 and 1993.

C.f. Teller (1983, p. 148): “By setting the truths of kind \( P \) as the physical truths and the truths of kind \( S \) as all other truths [we can] express materialism as the view that everything supervenes on the physical, that once the physical is set, the rest is already determined”.

Thanks to an editor for calling this quotation to my attention.

Broad’s reference to failure of prediction suggests an epistemological account of nothing over and aboveness, but in fact he took such failures to reflect a metaphysical distinction
between laws (as in the previous quotation). As McLaughlin (1992, p. 73) notes, “Emergentists often speak of emergent properties and laws as unpredictable from what they emerge from. But [ . . . ] the Emergentists do not maintain that something is an emergent because it is unpredictable. Rather, they maintain that something can be unpredictable because it is an emergent”.

18 Besides van Cleve 1990, see Chalmers 1996, Kirk 1996, and Stoljar 2000. On a clarificatory note, Chalmers’s use of supervenience to formulate physicalism departs from the usual motivation for doing so—namely, that supervenience does not entail reduction; for he takes supervening with “broadly logical necessity” to require something like functional or causal reduction of supervenient to base properties. I treat Chalmers’s formulation in §2.4; meanwhile, I continue to assume that supervenience connections are characterized by property correlations alone.

19 Two classic position pieces are Shoemaker 1980 and Swoyer 1982; for recent defenses and refinements see Elder 1994, Shoemaker 1998, and Ellis 2001. I won’t repeat the arguments of these philosophers here; below I discuss my reasons for finding this thesis plausible and well-motivated.


21 Proponents of this view sometimes qualify it: laws constraining what may exist in the same position—e.g., laws ruling out something’s having both positive and negative charge—may not be very different. I’ll take this restriction for granted in what follows.

22 See Hume 1739, especially Book I, Parts I and III. In accordance with his strict empiricism, Hume took the allowable forms of inference to be (a) deductive (as in mathematical inference) or (b) a matter of psychological association of ideas (born of impressions of the external or internal worlds), in accordance with the “uniting principles” of resemblance, spatiotemporal contiguity, and cause and effect (and where the third unifying principle was analyzed in terms of constant conjunctions involving the first two principles).

23 How do Lewis and Armstrong, self-professed naturalists, miss this fact about the Contingency view? Answer: they each work with non-standard definitions of naturalism. On Lewis’s understanding (p.c.), to be naturalist is to be non-supernaturalist; on Armstrong’s understanding (1989, p. 3), to be naturalist is to believe that only entities that are part of the actual space-time world exist. The Contingency view is not in tension with these non-standard understandings.

24 Such an account of over and aboveness will also violate this criterion by rendering physicalism compatible with Malbranchian occasionalism; for an occasionalist may maintain that the nature of God (or of whatever supernatural deity is at issue) is such that simplicity, consistency, or some other virtue demands that they bring about the same higher-order properties in every possible world where the relevant physical properties are instanced.

25 Thanks to a referee for pressing this point.

26 I’ll speak of ‘laws governing fundamental interactions’ as shorthand for ‘laws governing the entities associated with fundamental interactions’.

27 Thanks to a referee for raising this issue.

28 Since this support is grounded in the character of the actual laws, I don’t suppose that Holism is metaphysically necessary; but I do suppose that those general features of laws grounding the present support of Holism are reasonably taken to extend to laws governing emergent properties.

29 Kirk (1996, p. 250–1) objects to using weak supervenience to formulate physicalism on these grounds: “This is widely agreed to be unsatisfactory as a basis for stating physicalism. For it requires only that there be some physical property correlated with the mental one, not that the latter involve only the physical”. As I have been arguing, pace Kirk and others, stronger varieties of supervenience don’t guarantee that a supervening property “involve only the physical”, either.

30 As a referee pointed out, these restrictions have usually been imposed because a given supervenience thesis is too strong, as opposed to too weak, for purposes of formulating physicalism. For example, the falsity of a global supervenience thesis is consistent with the
(contingent) truth of physicalism. From this perspective, what is wanted is a weaker supervenience thesis—one required to hold only in possible worlds with the restricted supervenience base. True; but it is also true that such restrictions are aimed at ensuring that in the relevant worlds, physically unacceptable entities are absent; and from this (present) perspective what is wanted is a stronger supervenience thesis.

31 Post (1987) endorses a similar view: “Roughly, [the physically possible worlds] are the worlds in which the laws of physics are true (and in which the only entities are those [identical with something or other among the mathematical-physical entities])”.

32 Broad’s criterion fails to distinguish laws as physically acceptable or emergent, for there are uncontroversially physically acceptable phenomena (of which he wasn’t aware) that are reasonably taken to be in-principle unpredictable, including properties of chaotic systems (see Newman 1996 and Bedau 1997).

33 To be sure, in giving these examples I did not cite any specific uncontroversially emergent property; but this is to be expected, insofar as whether there are any such properties is still an open empirical question (and, moreover, if one could cite such a property, the physicalism debate would presumably already have been resolved). Citing a specific emergent property isn’t needed, of course, to establish the point that if there are any emergent properties, then there are plausible grounds for thinking that they are necessarily connected to the physical properties upon which they depend.

34 This point doesn’t tell against Horgan’s self-identification as a non-reductive physicalist, since he doesn’t personally advocate formulating physicalism in terms of superdupervenience.

35 The trajectory of this point (arrived at independently) is very similar to that traced in section I of Melnyk 1999.

References


