ABSTRACT

Supervenience physicalism attempts to combine non-reductionism about properties with a physical determination thesis in such a way as to ensure physicalism. I argue that this attempt is unsuccessful: the specific supervenience relation in question is either strong enough to ensure reductionism, as in the case of strong supervenience, or too weak to yield physical determination, as in the case of global supervenience. The argument develops in three stages. First, I propose a distinction between two types of reductionism, definitional and scientific, a distinction thanks to which we can reply to a standard objection against the ontological reductionism of strong supervenience. Second, I claim that because of “the problem of random distribution,” global supervenience needs strengthening to be an adequate relation to capture our physicalistic intuitions; and I show, in accordance with Stalnaker’s relevant proof, why a natural strengthening of global supervenience renders it equivalent to strong supervenience. Finally, I argue against Stalnaker about the possibility of a non-reductionist global supervenience. The upshot is that despite appearances, supervenience physicalism is a form of reductive physicalism.

Keywords: global supervenience, strong supervenience, physicalism, reductionism, properties

The definitive commitment of physicalism (or materialism) is a type of substance monism, the idea that there is only one type of substance and it is physical (or material).\(^1\) The physicalist claims that there is no such thing as a mental substance which is the seat of mental properties. The physicalist holds that the Cartesian argument from differences between physical and mental properties to the claim that the type of substance they are instantiated in must be different is mistaken: there is no difference between physical properties like having \(10^9\) neurons and being \(\text{H}_2\text{O}\) and mental properties like thinking, wondering, and believing with respect to the type of substance they are properties of.

For properties, however, physicalism appears to be more tolerant with respect to the range of alternatives it admits. The once famous option is reductive physicalism (or property monism), according to which all properties are physical properties and, taking facts as property-instantiations, all facts are physical facts. Reductive physicalism is generally considered in the way E. Nagel (1961) introduces it: a semantic thesis as to relations between the vocabularies of scientific theories or linguistic frameworks. On Nagel’s account, a theory \(T_1\) is reduced to a theory

\(^1\) I shall use “physical/physicalism” and “material/materialism” interchangeably.
Let us call the type of physicalism which holds that psychological theories are reducible to physical theories in Nagel's sense of the term **scientific reductionism**. If scientific reductionism is true, then there are no non-physical mental properties.\(^3\)

It is fair to say that reductive physicalism is almost entirely out of philosophical fashion. The majority of philosophers with physicalist aspirations now accept that there are irreducibly mental properties and define their physicalism in terms of a supervenience thesis: there can be no mental difference without a physical difference or, equivalently, it is necessary that if physical facts are fixed then mental facts are fixed.\(^4,5\) The central motivation behind supervenience physicalism is to specify a *minimal* physicalistic determination thesis. It is generally held that it is possible to capture the generic physicalistic intuition by the idea that there is a sort of determination relation between physical and mental properties and yet without further committing oneself to the strong thesis that mental properties are reducible to physical properties. Supervenience physicalism operates on a somewhat elusive ground: it purports to combine property dualism with a determination thesis in such a way as to ensure physicalism.\(^6\)

I agree that scientific reductionism poses an unnecessarily strong requirement for a respectable physicalist stance, and also that supervenience physicalism exemplifies a way of being a physicalist without being a scientific reductionist. However, I will argue that the standard supervenience relations in the literature do not succeed in combining property dualism with the determination thesis: the specific supervenience relation in question is either strong enough to ensure reductive physicalism, as in the case of strong supervenience, or too weak to yield determination, as in the case of global supervenience. The argument developed here is in the spirit of Kim's (1989)
claim that non-reductive physicalism is “a myth”, though we will see that the way Kim responds to an important objection is unsatisfactory because of a failure to distinguish definitional reductionism from scientific reductionism. Moreover, as Stalnaker (1996) proves, in at least one of the plausible ways to strengthen global supervenience to make it yield determination, global supervenience turns out to be equivalent to strong supervenience. Stalnaker argues that the proved equivalence between global and strong supervenience does not show that one holding global supervenience is thereby committed to the reductionism of strong supervenience. However, I will show that Stalnaker’s reasons for the possibility of a non-reductionist global supervenience aren’t compelling. The upshot is that despite appearances, supervenience physicalism is a form of reductive physicalism.

I.

Definitional reductionism is scientific reductionism minus the claim as to the deducibility of laws of a given high level theory from those of a given low level theory. Definitional reductionism claims, in particular, that mental predicates or properties are definable by physical predicates or properties; but it refrains from making commitments as to the prospects of the logical deduction of psychological laws.7

Definitional reductionism is weaker than scientific reductionism. That all the predicates of mental vocabulary are defined in terms of those of physical vocabulary need not guarantee scientific reductionism. This is because, in typical versions of scientific reductionism, for every predicate of mental vocabulary there must be a kind predicate of physical vocabulary, and definitional reductions might not satisfy this condition. Definitional reductions are free to use Boolean predicate-forming operations (complementation, conjunction, disjunction, and perhaps others like infinite conjunction and disjunction) to generate new predicates from the predicates of a certain reducing theory, while it is far from clear that predicates formed by such operations will always in turn be the kind predicates of the reducing theory, that is, predicates which are eligible to appear in the law-statements of the reducing theory. Ergo definitional reductions of mental properties do not imply the deducibility of psychological laws. It is not an exaggeration to claim that the current non-reductionist consensus depends on a strong emphasis upon this point.8

What I claim above is this: if the Fodorian constraints on scientific laws are acceptable, then we can have definitional reductionism without scientific reductionism. These Fodorian constraints render the second condition in our original definition of scientific reductionism significant. In this paper, I will not question the plausibility of these constraints. What I will try to do is, among other things, to show that these constraints and property monism can be simultaneously accepted. As we shall see in the following section, there are some philosophers (e.g., Kim) who (implicitly) take for granted that

7 I assume in what follows that the necessary coextensionality of A and B suffices for the definitional reduction of A to B. This is controversial but not harmful to the current discussion. Non-reductive supervenience physicalists think that the supervenience of A on B doesn’t even imply accidental coextensionality (see Hellman and Thompson 1975). So, for the purposes of the paper, necessary coextensionality is reduction enough.

8 The locus classicus of this non-reductionist theme is Fodor’s “Special Sciences” (1974).
the Fodorian constraints present a serious obstacle to property monism. I will argue that this is not the case.

Definitional reductionism is a form of ontologically reductive physicalism: it claims that mental predicates (formal mode) or properties (material mode) are definable in terms of physical predicates or properties. Whatever the mode of talk one chooses, definability implies ontological reduction. If definitional reductionism is true, then there are no irreducibly mental properties or no mental predicates the referents of which are not the referents of any physical predicates.

It is worth stressing that definitional reductionism is a claim about properties referred to by predicates but not a claim about concepts associated with them. So, it does not try to capture all aspects of meaning and thus need not be modelled on analytic (meaning) definitions like “the predicate ‘bachelor’ means what ‘unmarried man’ means”. Definitional reductions only try to get the references of predicates – that is, properties – right. When the property water is defined in terms of the properties of being H and being O, the result is that the property water is not a property over and above the property H₂O but not necessarily that the concept water is thereby reduced to the concept H₂O. Similarly, if mental properties are reducible to physical properties, then what follows is property monism but not necessarily conceptual monism.

The bottom line is that supervenience physicalism, in virtue of its opposition to property monism, is entitled to reject not only scientific reductionism but also definitional reductionism. A supervenience physicalist who thinks that since he rejects scientific reductionism he automatically rejects reductive physicalism is mistaken. There is surely a sense in which the view of a philosopher granting definitional but not scientific reductionism is non-reductionist, but the point is that this is not the sense according to which a supervenience physicalist can define his position. The basic opponent of supervenience physicalism is not scientific reductionism but definitional reductionism.

II.

One of Kim’s contributions to the supervenience literature is to clearly distinguish global supervenience from strong supervenience: the first one is a claim about whole worlds, while the second one is a claim about particular objects in worlds. Given that A and B are sets of properties, strong supervenience can be defined as follows:

A strongly supervenes on B iff for any worlds w and w’, and for any objects x and

---

9 Surely, if we take “definitional reductionism” in this way, the claim in the previous paragraph (“definitional reductionism is a form of ontologically reductive physicalism”) trivially follows. Take a physical predicate PH and a psychological predicate PS. According to definitional reductionism, PS is reduced to PH if they refer to the same property P. So, reducing a predicate to another predicate is a way of showing that there is only one property in question. The connection between the linguistic and the ontological domains is already contained within my definition of definitional reductionism. As an anonymous referee suggested, definitional reductionism as I take it can also be suitably (and, perhaps, less misleadingly) called “ontological reductionism”.

10 From now on, I will use words in italics for concepts and words in bold for properties.

11 I shall not discuss here what Kim calls “weak supervenience” because it is almost universally accepted that it is a too weak a determination relation for a robust physicalism.
x’, if x in w is B-indiscernible from x’ in w’, then x in w is A-indiscernible from x’ in w’.

The strong supervenience of mental properties on physical properties is the claim that any two objects which have the same physical properties have the same mental properties.

In “Concepts of Supervenience”, Kim has proved that strong supervenience entails a type of reductive physicalism: if A strongly supervenes on B, then any property in A is necessarily coextensive with a property in B’, where B’ is the set closed under the closure of Boolean property-forming operations on B-properties. The subvenient B’-property with which an A-property is necessarily coextensive will sometimes be infinitely disjunctive, and the disjuncts will be B-maximal properties defined as the strongest consistent properties constructed by Boolean operations on the B-properties.12

Kim’s proof that strong supervenience entails necessary biconditionals faces the following objection:

Some might dispute this line of thought on the ground that “nomological properties”, i.e., those that are admissible in laws, are not closed under Boolean operations –that is, these operations, when applied to such properties, do not always yield properties fit to appear in laws. B-maximal properties, their infinite disjunctions, and the like are “too complex”, “too artificial and unnatural”, and “too heterogeneous”, it is argued, to be “natural kinds”. (Kim 1984, 171-2)

The core of Kim’s reply rests on a distinction between predicates and properties:

A long Boolean combination of predicates would normally be complex qua predicate; on the other hand, the property it expresses need not inherit that complexity (the Boolean expression may be equivalent to a short and simple one) … [There] is no direct inference from the constructional details of properties to their complexity or artificiality, whatever these things may mean for properties. (Kim 1984, 172)

What Kim says here is that the complexity of the Boolean operations that we employ to construct new properties from a set of properties need not be inherited by those new properties: the predicate expressing that new property might be complex and perhaps infinite, but there is no reason to suppose that the same also goes for the property itself.

Kim’s reply invoking predicate/property distinction is not satisfactory. For one thing, Kim has not given any reason to believe that the complexity of predicates constructed by Boolean operations need not be inherited by the properties expressed by those predicates. And there is some reason to suppose that those properties will not be simple: the property F and G expressed by the complex predicate F and G constructed by Boolean

---

12 There is no question in the literature about the soundness of Kim’s proof and also of Stalnaker’s proof that will be mentioned in section IV. Here I take the two proofs for granted.
conjunction from the predicates \( F \) and \( G \) appears to be complex. One may argue that the property \( F \text{ and } G \) is indeed simple in some interesting sense of the term; but then one has to come up with a plausible account of why and how this is the case, and merely asserting that it is the case is far from enough. For another thing, even when we assume that the complexity of Boolean combinations of predicates is \textit{never} inherited by the new properties thus constructed, we still stand in need of a reason for believing that those new properties would be natural kind properties in addition to being simple. Notice that not every simple property is a natural kind property, that is, a property which is expressed by a predicate which is fit to appear in the law-statements of the reducing theory – e.g., \textit{being depressed} is a simple property but arguably not a natural kind.\(^{13}\)

A better reply to the objection above is that properties constructed by Boolean operations need \textit{not} be natural kind properties, if what is to be shown is that strong supervenience is a form of ontologically reductive physicalism. The objection to Kim’s proof presupposes that in order to show that strong supervenience is a form of reductive physicalism one is obliged to show that it entails scientific reductionism. However, specifying a form of reductive physicalism distinct from scientific reductionism, we have shown that this presupposition is false: showing that strong supervenience entails definitional reductionism is sufficient for showing that it is ontologically reductive. And, given Kim’s proof to the effect that strong supervenience entails necessary property-correlations, it turns out that strong supervenience is a form of reductive physicalism. This means that the objection at hand against Kim’s proof misfires.\(^{14}\)

III.

Kim’s proof that strong supervenience is a form of reductive physicalism leads most non-reductive physicalists to adopt global supervenience usually defined as follows:

\[ \text{A globally supervenes on B iff any B-indiscernible worlds are also A-indiscernible.} \]

Stalnaker (1996, 227) gives a perspicuous definition of the relevant notion of indiscernibility: “For any set of properties and relations B, two worlds \( w \) and \( w' \) are B-indiscernible iff there

\(^{13}\) I take \textit{being depressed} as a simple property because it is a property that is picked out by an (atomic) psychological predicate (i.e., \textit{depressed}), i.e., one which is not formed by any Boolean operations. And, I take it that it is arguably not a natural kind because I am inclined to think that there are possibly no universal laws about what some philosophers call “felt moods”. I agree that both claims are objectionable. One can reasonably ask: Is being depressed simple? Is it not a natural kind? I think these questions deserve a closer look, but I will not attempt it here. I would only like to point out that there can (at least \textit{prima facie}) be simple properties that are not natural kinds. If \textit{being depressed} is not good enough for your taste, you can replace it by another property.

\(^{14}\) If strong supervenience, as Kim’s proof shows, entails definitional reductionism, then one who holds strong supervenience is committed to property monism given that definitional reductionism implies property monism. This is what I claim here; a defender of strong supervenience cannot consistently be a property dualist. Now, what can a defender of strong supervenience say about psychological laws? This depends on what he thinks about the Fodorian constraints I mentioned above on laws. If he thinks that these constraints are plausible, then he may thus be committed to the idea that there are no laws at the supervenient level. This is because the base properties that are identical to the supervenient properties will surely be constructed by long Boolean combinations, and thus they will not be eligible to figure in law statements. But if he thinks otherwise, then he may still hold that there are psychological laws which are not reducible to physical laws. My point here is that what a defender of strong supervenience thinks about Fodorian constraints on laws has no effect on what his position implies about properties, which is property monism. Thanks for an anonymous referee for pressing this issue.
is a one-to-one correspondence between the domains of \( w \) and \( w' \), and any individual in the domain of \( w \) has the same B-properties as the corresponding individual from the domain of \( w' \) has in \( w' \)."\(^{15}\) This is the notion many philosophers have in mind when they talk about global supervenience.

However, global supervenience as it stands faces what I call “the problem of random distribution”: it does not block worlds in which mental properties are randomly distributed among their objects. It allows, for instance, worlds like \( w' \) which is physically and mentally exactly like the actual world \( w \) except for the fact that in \( w' \), my physical duplicate is my cat Mischa’s mental duplicate and Mischa’s physical duplicate is my mental duplicate. The worlds \( w \) and \( w' \) are not only physically indiscernible but also mentally indiscernible in the sense defined above. There is a one-to-one correspondence between the domains of \( w \) and \( w' \), and for any individual \( x \) in \( w \) there is an individual \( x' \) in \( w' \) which has the same mental properties \( x \) has in \( w \). Since any acceptable physicalism must avoid worlds like \( w' \), global supervenience at hand turns out to be too weak to yield the sort of determination relations definitive of physicalism.

A stronger notion of global supervenience is thereby needed. Global supervenience is to be redefined in terms of a notion of distributional sameness. What goes wrong with the case above is that the supervenient mental properties are not correctly distributed among the particular objects which instantiate the subvenient physical properties. Here is an informal definition of the strengthened global supervenience:

\[
A \text{ globally supervenes on } B \text{ iff any two worlds in which } B\text{-properties are distributed according to a certain pattern are worlds in which } A\text{-properties are distributed according to that pattern.}
\]

Stalnaker gives a formal treatment to this bare-bones definition:

\[
A \text{ globally supervenes on } B \text{ iff any two worlds that are } B\text{-indiscernible relative to a mapping from the domain of one onto the domain of the other are also } A\text{-indiscernible relative to the same mapping. (Stalnaker 1996, 228)}
\]

The notion of B-indiscernibility relativized to a mapping goes like this: two worlds \( w \) and \( w' \) are B-indiscernible relative to a function of the domain of \( w \) onto the domain of \( w' \) iff the function is one-to-one, and each individual in the domain of \( w \) has the same properties in \( w \) that the corresponding individual has in \( w' \).

Global supervenience thus defined implies that there can be no worlds which are physically exactly like our world but in which mental properties are randomly distributed among their objects. Strengthening the notion of global supervenience has thereby resulted in giving us a sort of determination relation preventing the possibility of certain worlds which must be blocked by physicalism. The question for the non-reductive physicalist is whether global supervenience at hand is weak enough not to yield definitional reductions. The now

\(^{15}\) McLaughlin (1997) and Sider (1999) call this notion of supervenience “weak global supervenience”, and the following notion to be discussed shortly “strong global supervenience”. For a careful discussion, see Bennett (2004).
strengthened global supervenience asserts that worlds with physically indiscernible objects are worlds the corresponding objects of which are mentally indiscernible. How does this substantially differ from strong supervenience which asserts that physically indiscernible objects are mentally indiscernible?

An instructive way to approach this question is via Kim’s (1989) argument that with regard to worlds physically discernible from the actual world w, global supervenience is compatible with certain anti-materialist scenarios. Kim claims, in particular, that global supervenience allows worlds like w’ which, though is physically exactly like w except for the location of a single atom, differs radically from w mentally. Notice that if, for global supervenience, the individuation of mental properties of objects in w were entirely holistic, that is, if the mental properties objects have were to depend on nothing but their causal and spatiotemporal relations to all other objects in w, then Kim’s objection would indeed be very powerful. A fully holistic global supervenience has nothing to say about how the mental properties in worlds like w’ are to be distributed; moreover since, on this conception of global supervenience, any minute change in the physical properties of objects can result in drastic changes in their mental properties, the constancy in human beings’ having mental properties in the actual world would be miraculous. And, since miracles are far cries from physicalism, “global determination without local determinations” cannot be the right slogan for global supervenience.

An attractive reply to Kim’s objection is that of Paull and Sider (1992), which marries the global nature of the distribution of mental properties among objects in global supervenience with an account of intrinsic and extrinsic mental properties. On this account, properties can be intrinsic to individuals or portions of worlds. Intrinsic mental properties supervene on the intrinsic physical characteristics of the individuals they are properties of. Moreover, an extrinsic property of an individual can be an intrinsic property of the portion of the world he belongs to. With these distinctions in mind, consider the world w’ again. Either the mental properties we have are intrinsic to some portion of the w-minus-the-lone-atom (i.e., don’t depend on that atom), in which case the possibility of radical mental difference in w’ is incompatible with global supervenience. Or, the mental properties we have are not intrinsic to any portion of the w-minus-the-lone-atom (i.e., depend on that atom), in which case the possibility of radical mental difference is to be granted. But this has no implications about the question whether global supervenience is too weak a determination relation: all that shows is that mentality is more bizarre than we initially thought.

Paull and Sider’s account makes it explicit that strong and global supervenience are equivalent when restricted to intrinsic mental properties: global supervenience, just like its strong cousin, tells that two physically indiscernible objects must have the same intrinsic mental properties in all possible worlds. Hence, concerning intrinsic properties, global and strong supervenience are equally reductive.

It is sometimes argued that strong supervenience cannot account for the extrinsic character of certain mental states. Since I and my doppelganger in Twin-Earth have by hypothesis

---

16 Horgan’s (1993) reply is very similar to the one developed by Paull and Sider.
the same intrinsic physical properties, strong supervenience appears to predict that our beliefs about the watery stuff in our worlds are the same, while those beliefs are indeed different: I have beliefs about H$_2$O while he has beliefs about XYZ. But there is little reason to restrict the physical properties of individuals to their intrinsic physical properties; and, if extrinsic physical properties are allowed into the subvenient set, strong supervenience predicts the differences between extrinsic properties of intrinsically identical individuals. I and my doppelganger are indiscernible with respect to our intrinsic physical properties, but we are not physically indiscernible given our differences in extrinsic physical properties. Notice that strong and global supervenience are equivalent when extrinsic properties are allowed into the subvenient physical set: the claim that extrinsic properties of individuals in w globally supervene on the distribution of intrinsic properties over objects in w makes the same assertion as the claim that they strongly supervene on the distribution of physical properties including extrinsic ones.

IV.

Stalnaker (1996, 238) has proved that the strengthened global supervenience of A on B is equivalent to the strong supervenience of A on B', where B' is a certain closure of the set of B-properties. In light of the above discussion, this is to be expected. The equivalence of global and strong supervenience is bad news for the non-reductive physicalist because it implies that if A globally supervenes on B, then every A-property is necessarily equivalent to a property definable in terms of B-properties.

Matters get complicated, however, because of the way Stalnaker himself construes the consequences of his equivalence proof. Stalnaker argues that even if global supervenience is equivalent to strong supervenience, one can still consistently hold the former one while denying the latter:

Consider a philosopher who accepts this global supervenience thesis but rejects the corresponding strong supervenience thesis. If this philosopher is consistent, it must be that he holds that the set of physical properties is not closed under definability. There are two reasons he might think this: (1) perhaps he thinks that certain properties definable in terms of physical properties are not themselves physical properties; or (2) perhaps it is because he has a robust conception of property according to which certain well defined attributes (if I may use the word “attribute” neutrally for any way of picking out a set of individuals) do not correspond to properties of any kind, physical or not. (Stalnaker 1996, 228)

Neither of the reasons presented by Stalnaker for holding global supervenience without committing oneself to reductive physicalism is compelling. Let us first take the philosopher with the reason (1). This philosopher thinks that the property with which we end up while employing Boolean property-forming operations on certain physical B-properties in an attempt to define a given A-property, say, F is not itself a physical property. But this philosopher is confused, if he takes this thought as a reason to deny the definability of mental properties in terms of physical properties: granting that the property that we end up with in such a definition procedure is not a physical property does not imply that F is
not definable in terms of physical B-properties. This is indeed something Stalnaker himself acknowledges: “The philosopher with the first reason might still say that all properties are definable in terms of physical properties” (Stalnaker 1996, 228). But the very problem is whether mental properties are definable in terms of physical ones, not whether the properties constructed by Boolean operations on physical properties are themselves physical. If our philosopher admits what Stalnaker says, then he admits nothing less than the reductionism of strong supervenience.18

Let us take the philosopher with the reason (2). This philosopher claims that the attribute with which we end up while employing Boolean property-forming operations on certain physical B-properties in an attempt to define F is not itself a property. However, what we say above *mutatis mutandis* applies to this philosopher: this philosopher is confused, if he takes this claim as a reason to deny the definability of mental properties in terms of physical properties. This is again because granting that the attribute that we end up with in such a definition procedure is not a property does not imply that F is not definable in terms of physical B-properties: what such a grant would imply is only that F is a physical attribute, which is itself definable in terms of physical properties.

V.

Let me finish by highlighting the central points that have emerged in our discussion. I have first identified a type of reductive physicalism - definitional reductionism - which is weaker than scientific reductionism. Kim’s proof that strong supervenience entails that mental properties are necessarily coextensive with some physical properties faces the objection that physical properties necessarily coextensive with mental properties will be “too complex” to yield scientific reductions. Relying on the distinction between definitional and scientific reductionism, I have shown that this objection misfires. Moreover, I have argued that for global supervenience to be an adequate determination relation to define physicalism, it needs strengthening. Yet, a natural improvement of global supervenience makes it equivalent to strong supervenience. I have then argued against Stalnaker that the equivalence of global supervenience to strong supervenience commits one who holds the former to the reductionism of the latter. The general result is that supervenience physicalism, strong or global, is a form of reductive physicalism.

**Acknowledgments:** I would like to thank Robert Brandom, Anil Gupta, and two anonymous referees for their helpful comments on earlier drafts.

---

18 How can a property that is definable in terms of physical properties fail to be a physical property? This is a question that needs to be addressed to Stalnaker’s imaginary philosopher. He may argue that a property defined in terms of physical properties can fail to be a natural kind physical property (thanks to an anonymous referee for this point), and say that natural kind properties are the only kind of properties that is of interest to him. Now, what I agree with this philosopher is that he may be right in saying that not all properties formed by Boolean operations on physical properties are natural kind physical properties. But, surely, I think that he is mistaken in thinking that all physical properties are natural kind physical properties. The set of physical properties (though, perhaps, not the set of natural kind physical properties) is closed under definability.
REFERENCES


Received: March 5, 2009
Accepted: March 30, 2011

Department of Philosophy
University of Pittsburgh
1001 Cathedral of Learning
Pittsburgh, PA 15206
USA
erd9@pitt.edu