8

Supervenience as a philosophical concept

I. SUPERVENIENCE IN PHILOSOPHY

Supervenience is a philosophical concept in more ways than one. First of all, like such concepts as cause and rationality, it is often used in the formulation of philosophical doctrines and arguments. Thus, we have the claim that ethical predicates are "supervenient predicates", or that the characteristics of a whole supervene on those of its parts. And arguments have been advanced to show that the supervenience of moral properties undermines moral realism, or that, on the contrary, moral supervenience shows ethical judgments are "objective" after all. And, again like causality and rationality, the concept of supervenience itself has become an object of philosophical analysis and a matter of some controversy.

But unlike causality, supervenience is almost exclusively a philosopher's concept, one not likely to be encountered outside philosophical dissertations and disputations. The notion of cause, on the other hand, is an integral part of our workaday language, a concept without which we could hardly manage in describing our experiences and observations, framing explanations of natural events, and assessing blame and praise. Something similar can be said about the notion of being rational as well, although this concept is not as ubiquitous in ordinary discourse as that of cause. Supervenience of course is not unique in being a technical philosophical concept; there are many others, such as "haecceity" and "possible world" in metaphysics, "analyticity" in the theory of meaning, and the currently prominent concepts of "wide" and "narrow" content.

But this isn't to say that the word "supervenience" is a philosopher's neologism; on the contrary, it has been around for some time, and has had a respectable history. The O.E.D. lists 1594 for the first documented

This is the Third Metaphilosophy Address, delivered in May 1989 at the Graduate School of the City University of New York.

occurrence of the adjective "supervenient" and 1647–48 for the verb "supervene"; the noun "supervenience" occurred as early as 1664. In these uses, however, "supervene" and its derivatives were almost without exception applied to concrete events and occurrences in the sense of "coming upon" a given event as something additional and extraneous (perhaps as something unexpected), or coming shortly after another occurrence, as in "Upon a sudden supervened the death of the king" (1647–48) and "The king was bruised by the pommel of his saddle; fever supervened, and the injury proved fatal" (1867). There is also this entry from Charlotte Brontë's *Shirley* (1849): "A bad harvest supervened. Distress reached its climax". In common usage supervenience usually implies temporal order: the supervenient event occurs after the event upon which it supervenes, often as an effect. It is clear that even though the vernacular meaning of "supervenience" is not entirely unrelated to its current philosophical sense, the relationship is pretty tenuous, and unlikely to provide any helpful guide for the philosophical discussion of the concept.

I noted that supervenience is like haecceity and narrow content in that they are specifically philosophical concepts, concepts introduced by philosophers for philosophical purposes. But in one significant respect supervenience differs from them: haecceity and narrow content are notions used within a restricted area of philosophy, to formulate distinctions concerning a specific domain of phenomena, or for the purpose of formulating doctrines and arguments concerning a specific topic. Thus, the notion of haecceity arises in connection with the problem of identity and the essence of things; and the concepts of narrow and broad content emerge in the discussion of some problems about meaning and propositional attitudes. In contrast, supervenience is not subject-specific. Although the idea of supervenience appears to have originated in moral theory,[1] it is a general, methodological concept in that it is entirely topic-neutral, and its use is not restricted to any particular problem or area of philosophy. It is this subject-neutral character of supervenience that distinguishes it from the usual run of philosophical concepts and makes it an appropriate object of metaphilosophical inquiry. Supervenience is a topic of interest from the point of view of philosophical methodology.

In undertaking a philosophical study of supervenience we quickly run into the following difficulty. Because the term is rarely used outside philosophy, there is not a body of well-established usage in ordinary or scien-

1 However, see below on emergence and footnote 5.

132

tific language that could generate reliable linguistic intuitions to guide the inquiry; there are few linguistic or conceptual data against which to test one's speculations and hypotheses. This means that for supervenience there are not the usual constraints on the "analysis" of a concept; in a sense, there is no pre-existing concept to be analyzed. As we shall see, earlier philosophical uses of the concept do set some broad constraints on our discussion; however, when it comes to matters of detail supervenience is going to be pretty much what we say it is. That is, within limits we are free to define it to suit the purposes on hand, and the primary measure of success for our definitions is their philosophical usefulness. This, I believe, is the principal explanation of the multiplicity of supervenience concepts currently on the scene.

Perhaps, the concept of a possible world is also like this. If we want to use this concept in a serious way, we would need to explain what we mean, either by explicitly defining it or by providing appropriately chosen examples and applications. However this is done, we need not be bound, in any significant way, by previous usage; there is not a common body of philosophical usage to which one's conception of a possible world must answer. The only criterion of success here is pragmatic: how useful and fruitful the introduced concept is in clarifying modal concepts, systematizing our modal intuitions, and helping us sharpen our metaphysical opinions.

There is a long tradition of philosophical discussion of modal concepts, the notions of necessity and possibility, of essential and contingent properties, of essences and haecceities, and so on. In contrast, supervenience is a concept of a comparatively recent origin. R. M. Hare is usually credited with having introduced the term "supervenience" into contemporary discussion, and our present use of the term appears historically continuous with Hare's use of it in *The Language of Morals* (1952).[2] More than thirty years later, in his Inaugural Address, "Supervenience" (1984), to the Aristotelian Society,[3] Hare wonders who first used the term in its current philosophical sense, being quite sure that he was not that person. Hare says that he first used the term in an unpublished paper written in 1950, but is not able to name any particular philosopher who had used it before he did. In any case, Hare's introduction of the term didn't exactly start a stampede. There were, to be sure, isolated appearances of the concept in

2 London: Oxford University Press, 1952.
3 R. M. Hare, "Supervenience", *The Aristotelian Society,* Supplementary Volume 58 (1984): 1–16.

133

the ethical literature during the two decades following the publication of *The Language of Morals*[4] but they were not marked by any real continuity, or an awareness of its potential and significance as a general philosophical concept. An idea related to supervenience, that of "universalizability" or "generality" of moral judgments, was much discussed in moral philosophy during this period, but the debate remained pretty much one of local concern within ethics.

It would be an error, however, to think that moral theorists had a monopoly on supervenience. On the contrary, early in this century, "supervenience" and its derivatives were used with some regularity by the emergentists, and their critics, in the formulation and discussion of the doctrine of "emergent evolution", and it seems possible that Hare and others got "supervenience", directly or indirectly, from the emergentist literature. Many of the leading emergentists were British (for example, G. H. Lewes, Samuel Alexander, C. Lloyd Morgan, C. D. Broad) and the emergence debate was robust and active in the 1930s and '40s. The doctrine of emergence, in brief, is the claim that when basic physicochemical processes achieve a certain level of complexity of an appropriate kind, genuinely novel characteristics, such as mentality, appear as "emergent" qualities. Lloyd Morgan, a central theoretician of the emergence school, appears to have used "supervenient" as an occasional stylistic variant of "emergent", although the latter remained the official term associated with the philosophical position, and the concept he intended with these terms seems surprisingly close to the supervenience concept current today.[5]

The emergence debate, however, has by and large been forgotten, and appears to have had negligible effects on the current debates in metaphysics, philosophy of mind, and philosophy of science, except perhaps in some areas of philosophy of biology.[6] This is to be regretted because some

4 In *Moral Notions* (London: Routledge & Kegan Paul, 1967), pp. 158–159, Julius Kovesi points to the same characteristic of "good" that Hare called supervenience, but without using the supervenience terminology. Kovesi mentions that this characteristic is had by many nonethical expressions as well, e.g., "tulip"; however, he does not develop this point in any detail.

5 See especially Morgan's *Emergent Evolution* (London: Williams and Norgate, 1923). Others who used "supervenience" in connection with the doctrine of emergence include Stephen C. Pepper, "Emergence", *Journal of Philosophy* 23 (1926): 241–45; and Paul Meehl and Wilfrid Sellars, "The Concept of Emergence", *Minnesota Studies in the Philosophy of Science*, vol. 1, ed. Herbert Feigl and Michael Scriven (Minneapolis: University of Minnesota Press, 1956).

6 See, e.g., Ernest Nagel, *The Structure of Science* (New York: Harcourt Brace & World, 1961); F. J. Ayala and T. Dobzhansky, eds., *Studies in the Philosophy of Biology: Reduction and Related Problems* (Berkeley and Los Angeles: University of California Press, 1974).

134

of the issues that were then discussed concerning the status of emergent qualities are highly relevant to the current debate on mental causation and the status of psychology in relation to the biological and physical sciences.[7] In any case, the present interest in supervenience was kindled by Donald Davidson in the early 1970s when he used the term in his influential and much discussed paper "Mental Events"[8] to formulate a version of nonreductive physicalism. What is noteworthy is that the term has since gained quick currency, especially in discussions of the mind–body problem; and, more remarkably, the term seems by now to have acquired, among philosophers, a pretty substantial shared content. "Supervene" and its derivatives are now regularly encountered in philosophical writings, and often they are used without explanation, signaling an assumption on the part of the writers that their meaning is a matter of common knowledge.

And during the past decade or so attempts have been made to sharpen our understanding of the concept itself. Various supervenience relations have been distinguished, their mutual relationships worked out, and their suitability for specific philosophical purposes scrutinized. This has led David Lewis to complain about an "unlovely proliferation" of supervenience concepts, which he believes has weakened its core meaning.[9] I disagree with the "unlovely" part of Lewis's characterization, but he is certainly right about the proliferation. I think this is a good time to take stock of the current state of the supervenience concept, and reflect on its usefulness as a philosophical concept. That is my aim in this essay.

II. COVARIANCE, DEPENDENCE, AND NONREDUCIBILITY

The first use of the term "supervene" (actually, the Latin "supervenire") I have found in a philosophical text is by Leibniz. In connection with his celebrated doctrine concerning relations, Leibniz wrote:

> Relation is an accident which is in multiple subjects; it is what results without any change made in the subjects but supervenes from them; it is the

7 I discuss the doctrine of emergence in relation to the currently popular doctrine of nonreductive physicalism in "'Downward Causation' in Emergentism and Nonreductive Physicalism", in *Emergence or Reduction?*, ed. A. Beckermann, H. Flohr, and J. Kim (Berlin: De Gruyter, 1992).

8 Reprinted in Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980); originally published in 1970.

9 *On the Plurality of Worlds* (Oxford & New York: Basil Blackwell, 1986), p. 14.

135

thinkability of objects together when we think of multiple things simultaneously.[10]

There has been much interpretive controversy concerning Leibniz's doctrine of relations – in particular, whether or not it was a reducibility thesis, to the effect that relations are reducible, in some sense, to "intrinsic denominations" of things. Leibniz's use of "supervene" in this context seems not inappropriate in our light: his thesis could be interpreted as the claim that relations supervene on the intrinsic properties of their relata. Such a claim would certainly be an interesting and important metaphysical thesis.

But Leibniz's use of "supervene" may well have been an isolated event; although I cannot say I have done anything like an exhaustive or systematic search, I have not found any other occurrence of the term since then, until we come well into the present century. However, the idea of supervenience, or something very close to it, if not the term "supervenience", was clearly present in the writings of the British Moralists. There is, for example, the following from Sidgwick:

> There seems, however, to be this difference between our conceptions of ethical and physical objectivity: that we commonly refuse to admit in the case of the former – what experience compels us to admit as regards the latter – variations for which we can discover no rational explanation. In the variety of coexistent physical facts we find an accidental or arbitrary element in which we have to acquiesce, . . . But within the range of our cognitions of right and wrong, it will generally be agreed that we cannot admit a similar unexplained variation. We cannot judge an action to be right for A and wrong for B, unless we can find in the nature or circumstances of the two some difference which we can regard as a reasonable ground for difference in their duties.[11]

Sidgwick is saying that moral characteristics must necessarily *covary* with certain (presumably nonmoral) characteristics, whereas there is no similar covariance requirement for physical properties. In terms of supervenience the idea comes to this: moral properties, in particular, the rightness or wrongness of an action, are supervenient on their nonmoral properties (which could provide reasons for the rightness or wrongness).

Concerning the concept of "intrinsic value", G. E. Moore said this:

---

10 *Die Leibniz-Handschriften der koenighlichen oeffentlichen Bibliothek zu Hannover,* ed. E. Bodemann, Hanover, 1895, VII, c, p. 74. Quoted by Hide Ishiguro in her *Leibniz's Philosophy of Logic and Language* (Ithaca: Cornell University Press, 1972), p. 71, fn. 3. The Latin text reads: "Relatio est accidens quod est in pluribus subjectis estque resultans tantum seu nulla mutatione facta ab iis supervenit, si plura simul cogitantur, est concogitabilitas".
11 *The Method of Ethics,* pp. 208–209. Quoted by Michael DePaul in his "Supervenience and Moral Dependence", *Philosophical Studies* 51 (1987): 425–439.

136

> . . . if a given thing possesses any kind of intrinsic value in a certain degree, then not only must that same thing possess it, under all circumstances, in the same degree, but also anything *exactly like it,* must, under all circumstances, possess it in exactly the same degree.[12]

Likeness of things is grounded, presumably, in their descriptive, or "naturalistic" properties – that is, their nonevaluative properties. Thus, Moore's point amounts to the statement that the intrinsic value of a thing supervenes on its descriptive, nonevaluative properties.

Hare, introducing the term "supervenience" into moral philosophy for the first time, said this:

> First, let us take that characteristic of "good" which has been called its supervenience. Suppose that we say, "St. Francis was a good man." It is logically impossible to say this and to maintain at the same time that there might have been another man placed exactly in the same circumstances as St. Francis, and who behaved in exactly the same way, but who differed from St. Francis in this respect only, that he was not a good man.[13]

It is clear that both Moore and Hare, like Sidgwick, focus on the characteristic of moral properties or ethical predicates that has to do with their *necessary covariation* with descriptive – nonmoral and nonevaluative – properties or predicates. The attribution of moral properties, or the ascription of ethical predicates, to an object is necessarily constrained, in a specific way, by the nonethical properties attributed to that object. For Moore the constraint has the modal force of "must"; for Hare, the violation of the constraint amounts to the contravention of logical consistency.

The basic idea of supervenience we find in Sidgwick, Moore, and Hare, therefore, has to do with property covariation: properties of one kind must covary with properties of another kind in a certain way. As Lewis put it, "no difference of one sort without differences of another sort";[14] and a change in respect of properties of one sort cannot occur unless accompanied by a change in respect of properties of another sort. If you have qualms about properties as entities, the same idea can be expressed in terms of predicates; if you think the predicates in question do not express properties, in something like the way ethical noncognitivists regard ethical predicates, you could express the idea in terms of "ascriptions" of predicates or the making of ethical judgments.[15]

---

12 *Philosophical Studies* (London, 1922), p. 261.
13 *The Language of Morals* (London, 1952), p. 145.
14 *On the Plurality of Worlds* (Oxford: Blackwell, 1986), p. 14.
15 See James Klagge, "Supervenience: Ontological and Ascriptive", *Australasian Journal of Philosophy* 66 (1988): 461–470.

137

Hare spoke of ethical and other evaluative predicates as "supervenient predicates", apparently taking supervenience as a *property* of expressions. But it is evident that the fundamental idea involves a *relation* between two sets of properties, or predicates, and that what Hare had in mind was the supervenience *of* ethical predicates *in relation to* nonethical, or naturalistic predicates. In fact, that was precisely the way Lloyd Morgan used the term, in the 1920s, some three decades before Hare; he used "supervenience" to denote a general relation, speaking of the supervenience of physical and chemical events "on spatiotemporal events",[16] and of deity as a quality that might be supervenient "on reflective consciousness".[17] As I said, Morgan used "supervene" and "emerge" as stylistic variants, and this means that supervenience is as much a general relation as emergence is.

Thus, Morgan and other emergentists were the first, as far as I know, to develop a generalized concept of supervenience as a relation, and their concept turns out to be strikingly similar to that in current use, especially in philosophy of mind. They held that the supervenient, or emergent, qualities necessarily manifest themselves when, and only when, appropriate conditions obtain at the more basic level; and some emergentists[18] took great pains to emphasize that the phenomenon of emergence is consistent with determinism. But in spite of that, the emergents are not reducible, or reductively explainable, in terms of their "basal" conditions. In formulating his emergentism, Morgan thought of himself as defending a reasonable naturalistic alternative to both mechanistic reductionism and such anti-naturalisms as vitalism and Cartesianism. Thus, Morgan's position bears an interesting similarity to the supervenience thesis Davidson has injected into philosophy of mind, and to many currently popular versions of nonreductive materialism which Davidson has helped inspire. In a passage that has become a bench mark to the writers on supervenience and nonreductive materialism, Davidson wrote:

> 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 alike in all

16 Morgan, *Emergent Evolution*, p. 9.
17 Morgan, *Emergent Evolution*, p. 30. I should add that Morgan was here expounding Samuel Alexander's doctrine of emergence, and that he is skeptical about these two supervenience theses.
18 See Arthur O. Lovejoy's distinction between "indeterminist" and "determinist" theories of emergent evolution in his "The Meaning of 'Emergence' and Its Modes", *Proceedings of the Sixth International Congress of Philosophy* (New York, 1927): 20–33.

physical respects but differing in some mental respects, or that an object cannot alter in some mental respects without altering in some physical respects. Dependence or supervenience of this kind does not entail reducibility through law or definition: if it did, we could reduce moral properties to descriptive, and this there is good reason to *believe* cannot be done. . . . [19]

Both Morgan and Davidson seem to be saying that mental phenomena are supervenient on physical phenomena and yet not reducible to them.

What Davidson says about the supervenience relation between mental and physical characteristics is entirely consonant with the idea of property covariation we saw in Sidgwick, Moore, and Hare. But he did more than echo the idea of the earlier writers: in this paragraph Davidson explicitly introduced two crucial new ideas, earlier adumbrated in the emergence literature, that were to change the complexion of the subsequent philosophical thinking about supervenience. First, supervenience is to be a relation of *dependence:* that which is supervenient is dependent on that on which it supervenes. Second, it is to be a *nonreductive* relation: supervenient dependency is not to entail the reducibility of the supervenient to its subvenient base.[20]

Davidson had his own reasons for attaching these two ideas to supervenience. The quoted paragraph occurs in his "Mental Events" just after he has advanced his "anomalous monism", the doctrine that mental events are identical to physical events even though there are no laws connecting mental and physical properties. In writing this passage, he is trying to mitigate the likely impression that anomalous monism permits no significant relationships between mental and physical attributes, positing two isolated, autonomous domains. His psychophysical anomalism, the thesis that there are no laws connecting the mental with the physical, has sundered the two domains; with the supervenience thesis he is trying to bring them back together. But not so close as to revive the hope, or threat, of psychophysical reductionism.

In any event, these two ideas, dependency and nonreductiveness, have become closely associated with supervenience. In particular, the idea that supervenience is a dependency relation has become firmly entrenched, so firmly that it has by now acquired the status of virtual analyticity. But I think it is useful to keep these three ideas separate; so let us summarize the three putative components, or desiderata, of supervenience:

---

19 "Mental Events", p. 214.
20 Note that "nonreductive" is also consistent with reducibility. Thus, "nonreductive" is to be understood as indicting a neutral, noncommittal position with regard to reducibility, not as an affirmation of irreducibility.

139

Covariance: Supervenient properties covary with their subvenient, or base, properties. In particular, indiscernibility in respect of the base properties entails indiscernibility in respect of the supervenient properties.

Dependency: Supervenient properties are dependent on, or are determined by, their base properties.

Nonreducibility: Supervenience is to be consistent with the irreducibility of supervenient properties to their base properties.

Obviously, covariance is the crucial component; any supervenience concept must include this condition in some form. The main issue, then, concerns the relationship between covariance and the other two components, and here there are two principal questions. First, can covariance yield dependence, or must dependence be considered an independent component of supervenience? Second, is there an interpretation of covariance that is strong enough to sustain supervenience as a dependency relation but weak enough not to imply reducibility? More broadly, there is this question: In what ways can these three desiderata be combined to yield coherent and philosophically interesting concepts of supervenience? I will not be offering definitive answers to these questions here; for I don't have the answers. What follows is a kind of interim report on the ongoing work by myself and others on these and related issues.

III. TYPES OF COVARIANCE

In the quoted passage above, Davidson writes as though he held that property covariation of the sort he is specifying between mental and physical properties *generated* a dependency relation between them. That is, mental properties are dependent on physical properties *in virtue of* the fact that the two sets of properties covary as indicated. Is this idea sound? But what precisely is covariance, to begin with?

It turns out that the simple statement of covariance in terms of indiscernibility has at least two distinct interpretations, one stronger than the other, depending on whether things chosen for comparison in respect of indiscernibility come exclusively from one possible world, or may come from different worlds. We can call them "weak" and "strong" covariance. Let A and B be two sets of properties, where we think of A as supervenient and B as subvenient. I state two definitions for each type of covariance:

140

Weak covariance I: No possible world contains things, x and y, such that x and y are indiscernible in respect of properties in B ("B-indiscernible") and yet discernible in respect of properties in A ("A-discernible").

Weak covariance II: Necessarily, if anything has property F in A, there exists a property G in B such that the thing has G, and everything that has G has F.

Strong covariance I: For any objects x and y and any worlds w_i and w_j, if x in w_i is B-indiscernible from y in w_j (that is, x has in w_i precisely those B-properties that y has in w_j), then x in w_i is A-indiscernible from y in w_j.

Strong covariance II: Necessarily, if anything has property F in A, there exists a property G in B such that the thing has G, and *necessarily* everything with G has F.

For both weak and strong covariance, the two versions are equivalent under certain assumptions concerning property composition.[21] However, it will be convenient to have both versions. The sole difference between strong covariance II and weak covariance II lies in the presence of the second modal expression "necessarily" in the former; this ensures that the G–F correlation holds across possible worlds and is not restricted to the given world under consideration. I have elsewhere called the two types of covariance "weak supervenience" and "strong supervenience" respectively; I am using the "covariance" terminology here since I am trying to keep the idea of covariance and that of dependence separate. This is a purely terminological decision; if we liked, we could continue to use the supervenience terminology here, and then raise the question concerning the relationship between supervenience and dependence.

How should we understand the modal term "necessarily", or quantification over possible worlds, that occurs in the statements of covariance? I believe that a general characterization of covariance, or supervenience, should leave this term as an unfixed parameter to be interpreted to suit specific supervenience claims. The standard options in this area include metaphysical, logico-mathematical, analytic, and nomological necessity.

Hare's and Davidson's original statements of supervenience seem neutral with respect to weak and strong covariance. Interestingly, however, both have since come out in favor of weak covariance. Hare for moral supervenience and Davidson for psychophysical supervenience. In his Inaugural Address "Supervenience", Hare says that "what I have always had

21 See my "Concepts of Supervenience", Essay 4 of this volume, and "'Strong' and 'Global' Supervenience Revisited", Essay 5 of this volume.

in mind is not what Kim now calls 'strong' supervenience. It is nearer to his 'weak' supervenience. . . ."[22]

Davidson has recently given an explicit account of the notion of supervenience that he says he had earlier in mind:

> The notion of supervenience, as I have used it, is best thought of as a relation between a predicate and a set of predicates in a language: a predicate p is supervenient on a set of predicates s if for every pair of objects such that p is true of one and not of the other there is a predicate of s that is true of one and not of the other.[23]

We can easily verify that this is equivalent to weak covariance II, of the unit set consisting of p on the set s.

Hare and Davidson are not alone in their preference for weak covariance. Simon Blackburn, who has used normative supervenience as a premise in his argument against moral realism, opts for weak covariance as his favored form of supervenience, at least for the case of moral properties.[24] On his account, if property F supervenes on a set G of properties, the following holds: in every possible world, if something has F, its total or maximal G-property, G*, is such that anything with G* has F. Blackburn stresses that this last universal conditional, "Everything with G* has F", is to be taken as a material conditional with no modal force,[25] which makes his concept exactly fit our weak covariance II.

IV. COVARIANCE AND DEPENDENCE

As may be recalled, Davidson has said that the mental is "supervenient, or dependent" on the physical; here he seems to be using "supervenient"

22 "Supervenience", p. 4. Hare's actual definition of supervenience, pp. 4–5, is a little difficult to interpret in terms of our present scheme, in part because he still does not explicitly relativize supervenience, treating "supervenient" as a one-place predicate of properties. But there is little question that his definition of "F is a supervenient property" comes to "F is weakly covariant with respect to (G, not-G)" (it isn't wholly clear to me whether G is to be thought of as existentially quantified, or contextually indicated).

23 In his "Replies to Essays X–XII" in *Essays on Davidson: Actions and Events,* ed. Bruce Vermazen and Merrill B. Hintikka (Oxford: Clarendon Press, 1985), p. 242.

24 See "Supervenience Revisited", *Exercises in Analysis,* ed. Ian Hacking (Cambridge University Press, 1985). His (S), on p. 49, corresponds to weak covariance II; his (?), on p. 50, to strong covariance II. His argument against moral realism depends on accepting (S), not (?), as the appropriate form of moral supervenience. In "The Supervenience Argument Against Moral Realism", *Southern Journal of Philosophy* 30 (1992): 13–38, James Dreier urges a reading of Blackburn's argument on which it is committed to strong covariance.

25 Blackburn's (S), which he takes to characterize his notion of supervenience, is a little more complicated; it contains the relational predicate "x underlies y". However, I be-

142

and "dependent" interchangeably, or perhaps the former as specifying a sense of the latter. We have just seen that it is weak covariance that he says he had in mind when he spoke of supervenience. So there is the following substantive question: Can weak covariance give us a sense of dependence? Or equivalently: Can weak covariance be a form of supervenience if supervenience is to be a dependency relation?

Weak covariance does place a constraint on the distribution of supervenient properties relative to the distribution of their base properties. The question is whether this constraint is strong enough to warrant our considering it a form of dependence or determination. As I have argued elsewhere,[26] the answer must be in the negative. For concreteness consider the weak covariance of mental on physical properties; this covariance is consistent with each of the following situations:

(1) In a world that is just like this one in the distribution of physical properties, no mentality is present.
(2) In a world that is just like this one in all physical details, unicellular organisms are all fully conscious, while no humans or other primates exhibit mentality.
(3) In a world that is just like this one in all physical details, everything exhibits mentality in the same degree and kind.

These are all possible under weak covariance because its constraint works only *within* a single world at a time: *the fact that mentality is distributed in a certain way in one world has absolutely no effect on how it might be distributed in another world.* Intra-world consistency of the distribution of mental properties relative to the distribution of physical properties is the only constraint imposed by weak covariance.

This evidently makes weak covariance unsuitable for any dependency thesis with modal or subjunctive force. And modal force is arguably a necessary aspect of any significant dependency claim. Thus, when we say that the mental is dependent on the physical, we would, I think, want to exclude each of the possibilities, (1)–(3).[27]

Not so with strong covariance: property-to-property connections between supervenient and subvenient properties carry over to other worlds. That is obvious from both versions of strong covariance. Consider version

lieve what he has in mind with (S) is best read, and restated, as a definition of "underlie", that is, the converse of "supervene".

26 E.g., in "Concepts of Supervenience".

27 For some interesting considerations in defense of weak covariance in connection with materialism, see William Seager, "Weak Supervenience and Materialism", *Philosophy and Phenomenological Research* 48 (1988): 697–709.

143

II: When applied to the psychophysical case, it says that if anything has a mental property M, then there is some physical property P such that the "P → M" conditional holds across all possible worlds. This supports in a straightforward way the assertion that the psychological character of a thing is entailed, or necessitated, by its physical nature. The strength of entailment, or necessitation, in this statement depends on how the modal term "necessarily" is interpreted, or alternatively, what possible worlds are involved in our quantification over them (e.g., whether we are talking about all possible worlds, or only physically or nomologically possible worlds, etc.).

But does strong covariance give us dependence or determination? If the mental strongly covaries with the physical, does this mean that the mental is dependent on, or determined by, the physical? As we saw, strong covariance is essentially a relation of entailment or necessitation. We notice this initial difference between necessitation and dependence: dependence, or determination, is usually understood to be asymmetric whereas entailment or necessitation is neither symmetric nor asymmetric. We sometimes speak of "mutual dependence" or "mutual determination"; however, when nonreductive physicalists appeal to supervenience as a way of expressing the dependence of the mental on the physical, they pretty clearly have in mind an asymmetric relation: they would say that their thesis automatically excludes the converse dependence of the physical on the mental. "Functional dependence", in the sense that the two state variables of a system are related by a mathematical function, may be neither symmetric nor asymmetric; however, what we want is *metaphysical* or *ontic* dependence or determination, not merely the fact that values of one variable are determined as a mathematical function of those of another variable.

It isn't difficult to think of cases in which strong covariance fails to be asymmetric: think of a domain of perfect spheres.[28] The surface area of each sphere strongly covaries with its volume, and conversely, the volume with the surface area. And we don't want to say either determines, or depends on, the other, in any sense of these terms that implies an asymmetry. There is only a functional determination, and dependence, both

---

28 This example is similar to the one used by Lawrence Lombard in his interesting and helpful discussion of covariance and dependence in *Events: A Metaphysical Study* (London: Routledge, Kegan Paul, 1986), pp. 225ff. My discussion here is indebted to Lombard, and also to Michael R. DePaul, "Supervenience and Moral Dependence", *Philosophical Studies* 51 (1987): 425–439; and Thomas R. Grimes, "The Myth of Supervenience", *Pacific Philosophical Quarterly* 69 (1988): 152–160.

ways; but we would hesitate to impute a metaphysical or ontological dependence either way.

Could we get a relation of dependency by requiring that the subvenient properties not also strongly covary with the supervenient properties? Let us consider the following proposal:[29]

A-properties depend on B-properties just in case A strongly covaries with B, but not conversely; that is, any B-indiscernible things are A-indiscernible but there are A-indiscernible things that are B-discernible.

In most cases of asymmetric dependence this condition appears to hold; for example, the mental strongly covaries with the physical, but the physical does not strongly covary with the mental; and similarly for the evaluative and the descriptive. Moreover, all of these examples involve large and comprehensive systems of properties. So the idea would be that when an asymmetric strong covariance obtains for two comprehensive systems of properties, a dependency relation may be imputed to them.

It isn't clear that this proposal states a necessary condition for dependence. For consider this: chemical kinds (e.g., water, gold, etc.) and their microphysical compositions (at least, at one level of description) seem to strongly covary with each other, and yet it is true, presumably, that natural kinds are asymmetrically dependent on microphysical structures. Here our mereological intuition, that macrophysical properties are asymmetrically dependent on microphysical structures, seems to be the major influence on our thinking, cancelling out the fact that the converse strong covariance may also be present. I admit that this is not a clear-cut example; for one thing, the converse strong covariance could perhaps be defeated by going to a deeper micro-level description; for another, one might argue that there is here no dependence either way, since being a certain chemical kind just *is* having a certain micro-structure.

It is even less clear whether the proposal states a sufficient condition for dependence. There is reason to think it does not. For what does the added second condition that B not covary with A really contribute? What is clear is this: the absence of strong covariance from B to A guarantees that B does not depend on A. For that means that there are objects with identical B-properties but with different A-properties. So the net effect of this added condition is just that B does not depend on A. The question

29 In "The Myth of Supervenience" Grimes considers a criterion of this form and rejects it as neither necessary nor sufficient. The possible counterexamples I consider below are consistent with Grimes's argument; however, only schematic examples are presented by Grimes.

145

then is this: Can we count on A to depend on B whenever A strongly covaries with B and B does not depend on A?

One might argue for an affirmative answer as follows: "Strong covariance between A and B requires an explanation, and it is highly likely that any explanation must appeal to an asymmetric relation of dependence. So either A depends on B or B depends on A; but the failure of strong covariance from B to A shows that B doesn't depend on A. Hence, A depends on B."

What this argument neglects, rather glaringly, is the possibility that an explanation of the covariance from A to B may be formulated in terms of a third set of properties. It seems clearly possible for there to be three sets of properties A, B, and C, such that A and B each depend on C, A covaries with B but B does not covary with A, and A does not depend on B.[30] Something like this could happen if, although both A and B covary with C, B makes finer discriminations than A, so that indiscernibility in regard to B-properties entails indiscernibility with respect to A, but not conversely.

As a possible example consider this: I've heard that there is a correlation between intelligence as measured by the IQ test and manual dexterity. It is possible that both manual dexterity and intelligence depend on certain genetic and developmental factors, and that intelligence strongly covaries with manual dexterity but not conversely. If such were the case, we would not consider intelligence to be dependent on, or determined by, manual dexterity.

Although the argument, therefore, has a serious flaw, it is not without value. Observed correlations of properties, especially between two comprehensive systems of properties, cry out for an explanation, and when no third set of properties is in the offing that might provide an appropriate ("common cause") explanation, it may be reasonable to posit a direct dependency relation between the two property families. The proposed criterion of dependent covariation says that if B fails to covary with A, that rules out the possibility that B depends on A, leaving A's dependency on B as the only remaining possibility. So the criterion may be of some use in certain situations; however, it cannot be regarded – at least, in its present form – as an "analysis" of supervenient dependence, since the needed further condition (i.e., that there not be a set C on which both A and B severally depend) itself makes use of the concept of dependence.

Trying to define dependence in terms of covariance is not likely to

---

30 Grimes makes a similar point in "The Myth of Supervenience", p. 157.

146

meet with complete and unambiguous success. Consider the case of causal dependence. Experience has taught us that we are not likely to succeed in defining an asymmetric relation of causal dependence, or causal directionality, in terms only of nomological covariations between properties or event kinds.[31] Unless, that is, we make a direct appeal to some relation that is explicitly asymmetric, like temporal precedence. We are not likely to do any better with supervenient dependence; the proposal above, with the further proviso that the strong covariation holds for two *comprehensive* sets of properties, may be close to the best that can be done to generate dependence out of covariation. All this points to the conclusion that the idea of dependence, whether causal or supervenient, is metaphysically deeper and richer than what can be captured by property covariance, even when the latter is supplemented with the usual modal notions.[32]

Much of the philosophical interest that supervenience has elicited lies in the hope that it is a relation of dependency; many philosophers saw in it the promise of a new type of dependency relation that seemed just right, neither too strong nor too weak, allowing us to navigate between reductionism and outright dualism. And it is the dependency aspect of supervenience, not the covariation component, that can sanction many of the usual philosophical implications drawn from, or associated with, supervenience theses concerning various subject matters. Often it is thought, and claimed, that a thing has a supervenient property *because,* or *in virtue of the fact that,* it has the corresponding base property, or that its having the relevant base property *explains* why it has the supervenient property. All these relations are essentially asymmetric, and are in the same generic family of relations that includes dependence and determination. Clearly, property covariation by itself does not warrant the use of "because", "in virtue of", etc., in describing the relationship any more than it warrants the attribution of dependence. Thus, if we want to promote the doctrine of psychophysical supervenience, intending it to include a claim of psychophysical dependence, we had better be prepared to produce an independent justification of the dependency claim which

---

31 For further discussion see J. L. Mackie, *The Cement of the Universe* (Oxford: Oxford University Press, 1974), ch. 7; David H. Sanford, "The Direction of Causation and the Direction of Conditionship", *Journal of Philosophy* 73 (1976): 193–207, and "The Direction of Causation and the Direction of Time", *Midwest Studies in Philosophy* 9 (1984): 53–75; Tom Beauchamp and Alexander Rosenberg, *Hume and the Problem of Causation* (New York and Oxford: Oxford University Press, 1981), ch. 6.

32 Could counterfactuals help? Perhaps; see, e.g., David Lewis, "Causation", *Journal of Philosophy* 70 (1973): 556–567; but also Grimes, "The Myth of Supervenience".

goes beyond the mere fact of covariance between mental and physical properties.

Property covariation *per se* is metaphysically neutral; dependence, and other such relations, suggest ontological and explanatory directionality – that upon which something depends is ontologically and explanatorily prior to, and more basic than, that which depends on it. In fact, we can think of the dependency relation as explaining or grounding property covariations: e.g., one might say that mental properties covary with physical properties because the former are dependent on the latter. Direct dependence, however, is not the only possible explanation; as we saw, two sets of properties may covary because each is dependent on a common third set.

The upshot, therefore, is this: it is best to separate the covariation element from the dependency element in the relation of supervenience. Our discussion shows that property covariation alone, even in the form of "strong asymmetric covariance", does not by itself give us dependency; in that sense, dependency is an additional component of supervenience. But the two components are not entirely independent; for it seems that the following is true: for there to be property dependence there must be property covariation. We can, therefore, distinguish between two forms of dependence, each based on one of the two covariation relations. Thus, "strong dependence" requires strong covariation, while "weak dependence" can do with weak covariation. What must be added to covariation to yield dependence is an interesting, and metaphysically deep, question. It's analogous, in certain ways, to J. L. Mackie's question as to what must be added to mere causal connectedness to generate "causal priority", or "causal directionality". Mackie and others have sought a single, uniform account of that in which causal priority consists; however, it isn't at all obvious that our question concerning dependence admits of a single answer. Evidently, dependency requires different explanations in different cases, and for any given case there can be competing accounts of why the dependency holds. Among the most important cases of supervenient dependence are instances of part–whole dependence ("mereological supervenience"), and these may constitute a special basic category of dependence. Concerning the supervenience of the moral on the naturalistic, the classic ethical naturalist will formulate an explanation in terms of meaning dependence or priority; the noncognitivist's account may involve considerations of the function of moral language and why its proper fulfillment requires consistency, in an appropriate sense, of moral avowals in relation to descriptive judgments. These cases seem fundamentally different from

148

one another metaphysically, and any "analysis" of dependence that applies to all varieties of dependence, I think, is unlikely to throw much light on the nature of dependence. We will briefly return to these issues in a later section.

V. COVARIANCE AND REDUCIBILITY

As previously noted, Davidson has been chiefly responsible for the close association of supervenience with both the idea of dependency and that of nonreducibility. Nonreducibility, however, has been less firmly associated with supervenience than dependency has been; and there has been some controversy as to whether supervenience is in fact a nonreductive relation. Also, it seems that the association of nonreducibility with supervenience has come about from the historical happenstance that Moore and Hare, who are well known for their supervenience thesis concerning the moral relative to the naturalistic, also formulated classic and influential arguments against ethical naturalism, the doctrine that the moral is definitionally reducible to the naturalistic.[33] So why not model a nonreductive psychophysical relation on supervenience? If the moral could be supervenient on the naturalistic without being reducible to it, couldn't the mental be supervenient on the physical without being reducible to it?[34] But it is possible that the sense of reduction Moore had in mind when he argued against the reducibility of the moral is very different from the concept of reducibility that is now current in philosophy of mind; and it is also possible that Moore was just mistaken in thinking that he could have supervenience without reducibility.

Moore's so-called "open question" argument suggests that the sort of naturalistic reduction he was trying to undermine is a *definitional* reductionism – the claim that ethical terms are analytically definable in naturalistic terms. Moreover, the argument is effective only against the claim that there is an *overt synonymy* relation between an ethical term and its purported naturalistic definition. For consider what the open question is intended to test: for any pair of expressions X and Y, we are supposed to determine whether "Is everything that is X also Y?" can be used to ask an intelligible and significant question. The idea is that if X is definable as Y (that is, if X and Y are synonymous), the question would not be an

33 There may also have been the influence of the emergentist doctrine that emergent properties are irreducible to their "basal" conditions.
34 For further discussion of supervenience in relation to nonreductive physicalism see my "The Myth of Nonreductive Materialism", Essay 14 of this volume.

intelligible one (consider: "Is everyone who is a bachelor also a male?" and "Does everything that is a cube also have twelve edges?"). It is clear that the logical equivalence of X and Y, or the fact that in some philosophical sense X can be "analyzed" as Y, etc., would not make the question necessarily unintelligible or lack significance. The nomological equivalence between X and Y probably was the furthest thing from Moore's mind; he pointedly says that even if we found a "physical equivalent" of the color yellow, certain "light-vibrations" as Moore puts it,[35] these light-vibrations are not what the term "yellow" *means*. So Moore's anti-naturalism was the denial of the definitional reducibility of ethical terms to naturalistic terms, where the notion of definition itself is extremely narrowly construed.

The kind of reduction Davidson had in mind in "Mental Events" is considerably wider than definitional reduction of the Moorean sort: the main focus of his antireductionist arguments is *nomological* reduction, reduction underwritten by contingent empirical laws correlating, and perhaps identifying, properties being reduced with those in the reduction base. Davidson's argument is two-pronged: the demise of logical behaviorism shows the unavailability of a definitional reduction of the mental, and his own psychophysical anomalism, the doctrine that there are no laws correlating mental with physical properties, shows that a nomological reduction isn't in the cards either.[36] Moore would have been unconcerned about nomological reducibility; his anti-naturalism apparently permitted strong, necessary synthetic a priori, relationships between the moral and the nonmoral.

I earlier noted that the issue of reducibility seemed less central to supervenience than that of dependence. It is somewhat ironic that covariance seems more intimately connected with reduction than it is with dependence. But before getting into the details we must know what we mean by reduction. Reduction is standardly understood as a relation between *theories,* where a theory is understood to consist of a distinctive theoretical vocabulary and a set of laws formulated in this vocabulary. The reduction of one theory to another is thought to be accomplished when the laws of the reduced theory are shown to be derivable from the laws of the reducer theory, with the help of "bridge principles" connecting terms of the reduced theory with those of the reducer.[37] Just what bridge laws are re-

35 *Principia Ethica,* p. 10.   36 For details see Davidson's "Mental Events".
37 This is the model of derivational reduction developed by Ernest Nagel in *The Structure of Science* (Harcourt, Brace & World, 1961). Whether this is the most appropriate model to be used in the present context could be debated; on this issue see William C. Wimsatt,

quired obviously depends on the strength of the two theories involved, and there seems very little that is both general and informative to say about this. The only requirement on the bridge laws that can be explicitly stated, independently of the particular theories involved, is the following, which I will call "the condition of strong connectibility":[38]

Each primitive predicate P of the theory being reduced is connected with a coextensive predicate Q of the reducer in a biconditional law of the form: "for all x, Px iff Qx"; and similarly for all relational predicates.

If this condition is met, then no matter what the content of the two theories may be, derivational reduction is guaranteed; for these biconditional laws would allow the rewriting of the laws of the theory being reduced as laws of the reducer, and if any of these rewrites is not derivable from the pre-existing laws of the reducer, it can be added as an additional law (assuming both theories to be true). In discussing reduction and covariance, therefore, we will focus on this condition of strong connectibility.[39]

To begin, weak covariance obviously does not entail strong connectibility. Weak covariance lacks an appropriate modal force to generate laws; as noted, the correlations entailed by weak covariance between supervenient and subvenient properties have no modal force, being restricted to particular worlds.

What then of strong covariance? Here the situation is different; for consider strong covariance II: it says that whenever a supervening property P is instantiated by an object, there is a subvenient property Q such that the instantiating object has it and the following conditional holds: necessarily if anything has Q, then it has P. So the picture we have is that for supervenient property P, there is a set of properties, Q_1, Q_2, . . . in the subvenient set such that each Q_i is necessarily sufficient for P. Assume that this list contains all the subvenient properties each of which is sufficient for P. Consider then their disjunction: Q_1 or Q_2 or . . . (or UQ_i, for short). This disjunction may be infinite; however, it

"Reductive Explanation: A Functional Account", in R. S. Cohen *et al.,* eds. *PSA 1974,* pp. 671–710.

38 Restricting ourselves to theories formulated in first-order languages.

39 There are various plausible considerations for thinking that derivational reduction as characterized isn't enough (and that it may not even be necessary). One line of consideration seems to show that we need *identities* of entities and properties rather than correlations; another line of consideration argues that the reduction must exhibit some underlying "mechanism", preferably at a micro-level, that explains how the higher processes work. We must bypass these issues here.

151

is a well-defined disjunction, as well-defined as the union of infinitely many sets. It is easy to see that this disjunction is necessarily coextensive with P.

First, it is clear enough that UQ_i entails P, since each disjunct does. Second, does P entail UQ_i? Suppose not: something then, say b, has P but not UQ_i. According to strong covariance, b has some property in the subvenient set, say S, such that necessarily whatever has S also has P. But then S must be one of the Q_i, and since b has S, b must have UQ_i. So P entails UQ_i. So P and UQ_i are necessarily coextensive, and whether the modality here is metaphysical, logical, or nomological, it should be strong enough to give us a serviceable "bridge law" for reduction.

So does this show that the strong connectibility is entailed by strong covariance, and hence that the supervenience relation incorporating strong covariance entails reducibility? Some philosophers will resist this inference.[40] Their concern will focus on the way the nomological coextension for P was constructed in the subvenient set − in particular, the fact that the constructional procedure made use of disjunction.[41] There are two questions, and only two as far as I can see, that can be raised here: (1) Is disjunction a proper way of forming properties out of properties? (2) Given that disjunction is a permissible property-forming operation, is it proper to form infinite disjunctions? I believe it is easy to answer (2): the answer has to be a yes. I don't see any special problem with an infinite procedure here, any more than in the case of forming infinite unions of sets or the addition of infinite series of numbers. We are not here talking about predicates, or linguistic expressions, but properties; I am not saying that we should accept predicates of infinite length, although I don't know if anything would go astray if we accepted infinite disjunctive predicates that are finitely specified (we could then introduce a simple predicate to abbreviate it). So the main question is (1).

Is disjunction a permissible mode of property composition? One might argue as follows for a negative answer, at least in the present context:

40 See Paul Teller, "Comments on Kim's Paper", and John Post, "Comment on Teller", both in *Southern Journal of Philosophy* 22 (1983), *The Spindel Conference Supplement* ("Supervenience"): 57–62, 163–167.

41 This in part meets an objection that John Post has raised (in his "Comment on Teller") against my earlier construction of these coextensions (in "Concepts of Supervenience") which made use of other property-forming operations. Post's specific objection was aimed at property complementation (or negation). On this issue see also William Seager, "Weak Supervenience and Materialism", and James Van Cleve, "Supervenience and Closure", *Philosophical Studies* 58 (1990): 225–238. Some remarks to follow in the main text are relevant to Post's point.

152

Bridge laws are laws and must connect nomological kinds or properties (so their predicates must be "lawlike", "projectible", and so on). However, from the fact that M and N are each nomic, it does not follow that their disjunction, M or N, is also nomic. Consequently, our constructional procedure fails to guarantee the nomologicality of the generated coextensions.

One might try to buttress this point by the following argument: the core concept of a property is *resemblance* – that is, the sharing of a property must ensure resemblance in some respect. We can now see that the disjunctive operation does not preserve this crucial feature of propertyhood (nor does complementation, one might add). Round objects resemble one another and so do red objects; but we cannot count on objects with the property of being *round or red* to resemble each other. This is why "conjunctive properties" present no difficulties, but "disjunctive properties", and also "negative properties", are problematic.

I do not find these arguments compelling. It isn't at all obvious that we must be bound by such a narrow and restrictive conception of what nomic properties, or properties in general, must be in the present context. When reduction is at issue, we are talking about theories, theories couched in their distinctive theoretical vocabularies. And it seems that we allow, and ought to allow, freedom to combine and recombine the basic theoretical predicates and functors by the usual logical and mathematical operations available in the underlying language, without checking each step with something like the resemblance criterion; that would work havoc with free and creative scientific theorizing. What, after all, is the point of having these logical operations on predicates? When we discuss the definitional reducibility of, say, ethical terms to naturalistic terms, it would be absurd to disallow definitions that make use of disjunctions, negations, and what have you; why should we deny ourselves the use of these operations in forming reductive bridges of other sorts? Moreover, it may well be that when an artificial-looking predicate proves useful, or essential, in a fecund and well-corroborated theory and gets entrenched, we will come to think of it as expressing a robust property, an important respect in which objects and events can resemble each other. In certain situations, that recognizing something as a genuine property would make reduction possible may itself be a compelling reason for doing so![42]

---

42 I wonder how "natural" the quantity $1/2\,(mv^2)$ looked before it was identified as kinetic energy.

153

Let me make a final point about this. The fact that for each supervenient property, a coextension – a qualitative coextension if not a certifiably nomic one – exists in the subvenient base properties means that there is at least the possibility of our developing a theory that will give a perspicuous theoretical description of this coextension, thus providing us with strong reason for taking the coextension as a nomic property. At least in this somewhat attenuated sense, strong covariance can be said to entail the possibility of reducing the supervenient to the subvenient. And we should note this: if we knew strong covariance to fail, that would scotch the idea of reduction once and for all.

We should briefly look at "global supervenience", or "global covariance", as a nonreductive supervenience relation. For this idea has been touted by many philosophers as an appropriate dependence relation between the mental and the physical which is free of reductive implications.[43] The basic idea of global supervenience is to apply the indiscernibility considerations globally to "worlds" taken as units of comparison. Standardly the idea is expressed as follows:

Worlds that are indiscernible in respect of subvenient properties are indiscernible in respect of supervenient properties.

Worlds that coincide in respect of truths involving subvenient properties coincide in respect of truths involving supervenient properties.

For our present purposes we may think of indiscernibility of worlds in respect of a given set of properties (say, physical properties) as consisting in the fact that these properties are distributed over their individuals in the same way (for simplicity we may assume that the worlds have the same individuals).

It is known that this covariance relation does not imply property-to-property correlations between supervenient and subvenient properties; thus, it does not imply what I have called strong connectibility.[44] So global supervenience, along with weak supervenience, can qualify as a nonreductive relation. But this is a signal that global covariance may be quite

---

43 See, e.g., Terence Horgan, "Supervenience and Microphysics", *Pacific Philosophical Quarterly* 63 (1982): 29–43; David Lewis, "New Work for a Theory of Universals", *Australasian Journal of Philosophy* 61 (1983): 343–377. Also Geoffrey Hellman and Frank Thompson,"Physicalism: Ontology, Determination, and Reduction", *Journal of Philosophy* 73 (1975): 551–564.

44 See, e.g., Bradford Petrie, "Global Supervenience and Reduction", *Philosophy and Phenomenological Research* 48 (1987): 119–130. (Added 1993: for further discussion of this issue, see "Postscripts on Supervenience", section 3, in this volume.)

154

weak, perhaps too weak to sustain a dependency relation of significance.[45]

As I have argued elsewhere,[46] this can be seen in at least two ways. First, this form of covariance permits worlds that differ minutely in subvenient properties to differ drastically in respect of supervenient properties. Thus, global covariance of the mental with respect to the physical is consistent with there being a world that differs from this world in some insignificant physical detail (say, it contains one more hydrogen atom) but which differs radically in psychological respects (say, it is wholly void of mentality). Second, global covariance as explained fails to imply weak covariance; that is, it can hold where weak covariance fails. This means that psychophysical global covariance can be true in a world that contains exact physical duplicates with divergent psychological characteristics; it permits the existence in the actual world of an exact physical replica of you who, however, has the mentality of a fruit fly. There certainly is reason to wonder whether a supervenience relation whose property covariance requirement is this weak can qualify as a dependency relation. As I argued earlier, property covariance alone, even "strong covariance", does not yield dependence, and in that sense dependence must be considered an independent component of supervenience in any case. However, again as I argued, dependence does require an appropriate relation of property covariance. This raises the following question: Is global covariance strong enough to ground a respectable supervenience relation? We may well wonder whether a supervenience relation based on global covariance might not turn out to be incongruous in that, given this is what it requires of property covariance, the dependency component makes little sense.

I suggest, however, that we keep an open mind about this, and adopt an attitude of "Let one hundred supervenience concepts bloom!" Each may have its own sphere of application, serving as a useful tool for formulating and evaluating philosophical doctrines of interest. And this does not mean that we must discard the core idea of supervenience captured by the maxim "No difference of one kind without a difference of another kind". It's just that we now recognize that this core idea can be explained in distinct but interestingly related ways, and that what we want to say about a supervenience claim about a specific subject matter may depend on the interpretation of supervenience appropriate to the context. I think this is philosophical progress.

45 I believe this indeed is the case with psychophysical global supervenience; for details see my "'Strong' and 'Global' Supervenience Revisited" and "The Myth of Nonreductive Materialism".
46 "'Strong' and 'Global' Supervenience Revisited".

155

VI. GROUNDS OF SUPERVENIENCE

It has been argued that supervenience is a mysterious and unexplained relation, and hence that any philosophical argument couched in the vocabulary of supervenience is a retrogressive and obfuscating maneuver incapable of yielding any illumination for the issue on hand. For example, Stephen Schiffer takes a dim view of those who appeal to supervenience:

> How could being told that non-natural moral properties stood in the supervenience relation to physical properties make them any more palatable? On the contrary, invoking a special primitive metaphysical relation of supervenience to explain how non-natural moral properties were related to physical properties was just to add mystery to mystery, to cover one obscurantist move with another. I therefore find it more than a little ironic, and puzzling, that supervenience is nowadays being heralded as a way of making nonpleonastic, irreducibly non-natural properties cohere with an acceptably naturalistic solution to the mind–body problem. . . . the appeal to a special primitive relation of 'supervenience', as defined above, is obscurantist in the extreme.[47]

The supervenience relation Schiffer refers to "as defined above" is in effect our strong covariance II, with the further proviso that the relationship "necessarily everything with G has F" is an unexplainable "brute metaphysical fact".

There perhaps have been philosophers who deserve Schiffer's excoriations; however, we need to separate Schiffer's editorial comment that supervenience is a "brute metaphysical" fact from a mere claim of supervenience concerning a given topic. Schiffer's addition is a nontrivial further claim, which someone advocating a supervenience thesis might or might not wish to make, that goes beyond the claim of supervenient covariance or dependence. For there is nothing in the concept of covariance or dependence that forces us to view supervenience as invariably involving unexplainable relationships. In fact, when a supervenience claim is made, it makes perfectly good sense to ask for an *explanation* of why the supervenience relation holds. Why does the moral supervene on the nonmoral? Why do facts about wholes supervene, if they do, on facts about their parts? Why does the mental supervene on the physical?

It may well be that the only answer we can muster for some of these questions is that, as far as we can tell, it is a brute fact. But that need not be the only kind of answer; we should, and can, hope to do better. This

---

47 *Remnants of Meaning* (Cambridge: MIT Press, 1987), pp. 153–154.

is evident from the following fact alone: supervenience, whether in the sense of covariation or in the sense that includes dependence, is transitive. This means that it is possible, at least in certain situations, to answer the question "Why does X supervene on Y?" by saying that, as it turns out, X supervenes on Z, and Z in turn supervenes on Y. The interpolation of another supervenient tier may well explain why X-to-Y supervenience holds. (Compare: Why does X cause Y? Answer: X causes Z, and Z causes Y.) As Schiffer says, Moore gave a sort of "brute fact" account of moral supervenience, and given his metaethical theory he probably had no other choice: we "intuit" necessary synthetic a priori connections between nonnatural moral properties and certain natural properties. But it isn't just ethical intuitionists like Moore who accept moral supervenience; Hare, whose metaethics radically diverges from Moore's, too has championed moral supervenience. And we also have Blackburn, a "projectivist" moral antirealist, who professes belief in moral supervenience, not to mention John Post,[48] who is an objectivist about ethical judgments. As I take it, these philosophers would give different accounts of why moral supervenience obtains; as we noted in our earlier discussion of dependence as a component of supervenience, Hare would presumably give an account in terms of some consistency requirement on the use of language of prescription. And Blackburn has argued against "moral realism" on the ground that it, unlike his own projectivist "quasi-realism", is unable to give a satisfactory explanation of moral supervenience.[49]

We may distinguish between two kinds of request for a "ground" of a supervenience relation. One concerns *general* claims of supervenience: why a given family of properties, say mental properties, supervene on another family, say neurobiological properties. Why does the mental supervene on the physical, and why does the normative supervene on the nonnormative? These are perfectly good, intelligible questions, which may or may not have informative answers. The second type of request concerns the relationship between *specific* supervenient properties and their base properties: Why is it that pain supervenes on the activation of A-delta and C-fibers? Why doesn't, say, itch or tickle supervene on it? Why doesn't pain supervene on, say, the excitation of A- and B-fibers?

The potential for supplying explanations for specific supervenience re-

48 See his "On the Determinacy of Valuation", *Philosophical Studies* 45 (1984): 315–333.
49 See Blackburn, "Supervenience Revisited". For discussion of Blackburn's argument see James Klagge, "An Alleged Difficulty Concerning Moral Properties", *Mind* 93 (1984): 370–380; James Dreier, "The Supervenience Argument Against Moral Realism".

lationships varies for different mind–body theories. Both the behaviorist and the functionalist could formulate a plausible meaning-based explanation (I mean, plausible *given* their basic doctrines): pain, not itch, supervenes on physicalistic condition P because of an analytic, semantic connection between "pain" and the standard expression for P. For the behaviorist, the connection is a direct one of definability. The functionalist will appeal to an additional empirical fact, saying something like this: "pain", as a matter of meaning, designates a certain causal-functional role, and it turns out, as a contingent empirical fact, that condition P occupies this causal role (in organisms or structures of a given kind). The functionalist can push ahead with his search for explanations and ask: why does condition P occupy this causal role in these organisms? This question is an empirical scientific question, and may be given an evolution-based answer, or one based in engineering considerations (in the case of artifacts); and there may be answers of other types.

There are philosophers who have a fundamentally physicalist outlook on the mind–body problem and yet would reject any analytic, definitional relationships between mental and nonmental expressions. Many of them would accept a thoroughgoing dependence of the mental on the physical grounded in lawlike type–type correlations between the two domains. Epiphenomenalism is such a position; so is the classic nonfunctionalist type-identity theory based on the supposed existence of pervasive psychophysical correlations. It seems that someone holding a physicalist position like these has no choice but to view the relationship between, say, pain and C-fiber activation as a brute fact that is not further explainable, something like the way G. E. Moore viewed the relationship between the nonnatural property of goodness and the natural property on which it supervenes. In this respect, the position of a physicalist who accepts psychophysical supervenience, especially of the "strong covariance" sort, but rejects a physicalist rendering of mental expressions, is much like that of those emergentists who regarded the phenomena of emergence as not susceptible of further explanation; that is, it is not further explainable why mentality emerges just when these physicochemical conditions are present, but not otherwise. Samuel Alexander, a leading emergentist, recommended that we accept these emergence relationships "with natural piety"; Lloyd Morgan, referring to Alexander, announced, "I accept this phrase".[50]

Is this a serious blemish on nonfunctionalist physicalism? This is an

50 *Emergent Evolution*, p. 36.

158

interesting, and difficult, question. Its proponents might insist that all of us must accept certain brute facts about this world, and that it is necessary to count fundamental psychophysical correlations among them in order to develop a plausible theory of mind, all things considered. This is only an opening move in what is likely to be a protracted dialectic between them and the functionalists, something we must set aside.[51] I will conclude with some brief remarks concerning explanations of general supervenience claims.

I think that the only direct way of explaining why a general supervenience relation holds, e.g., why the mental supervenes on the physical, is to appeal to the presence of specific supervenience relations – that is, appropriate correlations between specific supervenient properties and their subvenient bases. If these specific correlations are themselves explainable, so much the better; but whether or not they are, invoking them would constitute the first necessary step. Moreover, such correlations seem to be the best, and the most natural, *evidential ground* for supervenience claims – often the only kind of solid evidence we could have for *empirical* supervenience claims. Even the nonfunctionalist physicalist has an explanation of sorts for psychophysical supervenience: it holds because a pervasive system of lawlike psychophysical correlations holds. These correlations are logically contingent and empirically discovered; though they are not further explainable, they constitute our ground, both evidential and explanatory, of the supervenience of the mental on the physical.

This shows why a global supervenience claim *unaccompanied by the corresponding strong supervenience (or covariance) claim* can be so unsatisfying: we are being asked, it seems to me, to accept a sweeping claim about *all possible worlds,* say, that no two worlds could differ mentally without differing physically, on faith as a brute fact. In the absence of specific psychophysical correlations, and some knowledge of them, such a supervenience claim should strike us as a mere article of faith seriously lacking in motivation both evidentially and explanatorily; it would assert as a fact something that is apparently unexplainable and whose evidential status, moreover, is unclear and problematic. The attitude of the friends of global psychophysical supervenience is not unlike that of Samuel Alexander and Lloyd Morgan toward emergence: we must accept it "with natural piety"!

---

51 See for further discussion Terence Horgan and Mark Timmons, "Troubles on Moral Twin Earth: Moral Queerness Revisited"; *Synthese* 92 (1992): 221–260. Ernest Sosa has pointed out to me that appeals to meaning and analyticity, too, might involve appeals to brute facts in the end.

But there is this difference: the emergentists could at least point to the observed lawful correlations between specific mental and biological processes as evidence for the presence of a general system of such correlations encompassing all mental processes, and point to the latter as the ground of the general thesis of mental emergence.

160