

# Supervenience: Model Theory or Metaphysics?

JAMES C. KLAGGE

There are two problems concerning the implications of certain forms of supervenience that I wish to discuss. The problems are connected in that their resolution depends on deciding what sort of enterprise we are engaged in when we approach the problems. Supervenience can be formulated and discussed as a purely logical set of formulas, which are indeed quite engaging in their own right. On the other hand, much of the interest in supervenience has been generated by its apparent usefulness in understanding certain philosophically perplexing realms of life, for example, mentality and morality. These two conceptions of supervenience can come into conflict with one another, as we will see. The conflicts provide the opportunity to assess our motivations.

## 1. The First Problem

Suppose we accept supervenience in the form of the Quinean slogan "No difference without a physical difference," or, as Davidson puts it for one particular case, "There cannot be two events alike in all physical respects but differing in some mental respects."<sup>1</sup> Does it follow from this that there are principles, that is, universal generalizations, in which sufficient conditions are given in physical (or subvening) terms for the presence of certain supervening qualities?<sup>2</sup>

It might seem as though there obviously are such principles. Consider some mental state that I am now in, say one of anxiety. By our assumption, no one can be nonanxious without also differing from me in some physical way. In other words, if someone were exactly like me physically, that person would have to be anxious, too. Now, one may despair of ever giving a complete physical description of me, but nevertheless our assumption commits us to the principle that if someone has my physical description completely and exactly, that person is anxious. So, whatever my physical state is, call it  $P$ , it follows that

$$(\forall x)(Px \rightarrow x \text{ is anxious}).$$

Thus, the assumption of supervenience entails the existence of prin-

ciples that are one-directional and in which the antecedent states only physical conditions and the consequent attributes some supervening condition. As Richard Hare put it, "Supervenience brings with it the claim that there is some law [by which he seems to mean simply a universal generalization] which binds what supervenes to what it supervenes upon."<sup>3</sup>

Let us formalize the original assumption as

$$(1) \sim \Diamond (\exists x)(\exists y)[(\exists M)(Mx \ \& \ \sim My) \ \& \ (\forall P)(Px \equiv Py)]$$

(where  $M$  ranges over the supervening properties – e.g., mental properties – and  $P$  ranges over the subvening properties – e.g., physical properties). Or, more smoothly,

$$(2) \Box (\forall x)(\forall y)[(\forall P)(Px \equiv Py) \rightarrow (\forall M)(Mx \rightarrow My)].$$

And let us formalize the assertion of the existence of principles as

$$(3) \Box (\forall x)(\forall M)\{Mx \rightarrow (\exists P)[Px \ \& \ (\forall y)(Py \rightarrow My)]\}.$$

In other words, for any supervening condition, there is a sufficient condition that can be articulated in subvening terms. The subvening characterization may have to be an exhaustive, or maximal, characterization, as we supposed it was in giving the sufficient condition for my anxiety, but it may not be. Intuitively we can think of  $P$  in formula (3) as ranging over conjunctions of "mentally relevant" properties. But since we cannot generally delimit the set of mentally relevant properties in advance, we allow  $P$  to range over all (conjunctions of) physical properties.

Now we can formally state the first problem as whether (2) implies (3). (In the technical terminology of the literature, this is the problem of whether possible-worlds supervenience implies modal-operator supervenience.) I have already given an intuitive argument that it does, in which I was simply repeating a more formal proof offered by Jaegwon Kim.<sup>4</sup>

Kim's proof is impeccable from a model-theoretic point of view. However, the assumptions that lie behind this proof have been questioned by John Post.<sup>5</sup> Post objects as follows: the proof employs the notion of an exhaustive (or maximal) physical characterization of something. Presumably this characterization will indicate for each physical property whether or not the thing has the property.<sup>6</sup> Thus, the exhaustive characterization will be a conjunction of some physical properties and negations of the other physical properties. Post claims,<sup>7</sup> quite plausibly, that the negation of a physical property is not in general itself a physical property. For example, the property *not being an electron* is not a physical property, since it is had by the number 5, but the number 5 has no physical properties. Thus, a conjunction of physi-

cal properties and negations of physical properties is not in general a physical property. The assumption that the class of physical properties is closed under negation and conjunction, while admissible from a purely model-theoretic point of view, seems questionable after a metaphysical examination of the nature of physical properties. So the principle whose existence is asserted by formula (3) does not necessarily offer a sufficient condition *in physical terms* for the supervening property after all, and we cannot infer the existence of psychophysical principles from the assumption of psychophysical supervenience. It seems that metaphysics has to take precedence over model theory in our understanding of supervenience.

As a point about the metaphysical nature of the physical in general, I think Post's argument has to be accepted. The interest of the original proof is not mainly logical, but metaphysical – it purports to tell us something about the nature of the relationship between certain realms, such as the physical and the mental, or the natural and the moral. I agree that it is from a metaphysical perspective that we have to assess the assumptions that lie behind the proof. But Post's handling of the assumptions is still too abstract for the issue at hand.

Let us call the negations (or complements) of physical properties "quysical" properties. (That is short for "quasi-physical" properties.) Post's point is that the set of physical properties is not closed under complementation. Let us call a set of properties a "physical description" of an object if the set of properties

- (a) holds of the object in question,
- (b) includes only physical and quysical properties, and
- (c) includes at least one physical property.

Then it is intuitively true that, if supervenience is assumed, and an object with a set of *physical* properties has a supervening property, then there is a physical description (in my defined sense) such that any object that has that physical description will also have the supervening property in question.<sup>8</sup> So the psychophysical principles are back in business.

Is this an ad hoc trick? From the point of view of Post's metaphysical reflections on the nature of the physical, it may seem to be. But from the point of view of the metaphysics of mind and morality, I think it is not. What, then, are its motivations?

In morality it is not generally sufficient, when justifying a moral judgment (i.e., an all-things-considered judgment, not a *prima facie* judgment), to list some of a person's or an action's natural qualities – for a list of qualities may be "defeated" by the presence of some additional quality. For example, it may not be enough to explain Socrates' goodness to say he is honest, courageous, and wise if being bad-

tempered would cause us to rescind the judgment. So our moral principles may need to exclude some qualities as well as assert others. Until we have accounted for all the (potentially) morally relevant properties, our judgment must remain tentative (or *prima facie*). Thus, a principle can achieve some kind of closure only by asserting either the negation of every other morally relevant property or a negative existential to the effect that there are no other morally relevant properties.<sup>9</sup>

Perhaps, however, the search for closure (which is what led to the need for negations of physical properties) has more of a psychological than logical motivation. (If so, then we could ignore Post's metaphysical qualms about negations.) By calling a principle (or the resulting judgment) "*prima facie*," we mean that it holds so far as *these* considerations (covered in the antecedent of the principle) are concerned but may not hold once a wider range of considerations are brought to bear. By contrast, a principle (or resulting judgment) holds all things considered just insofar as all (relevant) considerations have been brought to bear. We cannot tell, just by examining a principle, whether it is (to be treated as) *prima facie* or not. Explicit indication of closure is a psychologically useful way to let us know that the principle (or the resulting judgment) is (meant as) all things considered. But it is not logically necessary. If being bad tempered would preclude Socrates from being good, then we should not assert:

$(\forall y)[(y \text{ is honest} \ \& \ y \text{ is courageous} \ \& \ y \text{ is wise}) \rightarrow y \text{ is good}]$ .

On the other hand, if we do assert this, then we would be logically committed to the irrelevance of Socrates' temper.

But this is too simplistic. Indication of closure – that a principle is meant as all things considered – is *not* just a psychological aid. In many contexts we could indicate closure by prefixing Frege's assertion sign ("⊢") to the principle. But what of other contexts in which contemplation or discussion of the principle is of interest or in which the principle is being assumed in the course of a derivation? In these contexts, since they are explicitly nonassertoric, we cannot rely on the assertion sign. If we want to draw out the implications of a principle without asserting it, we need to know and indicate whether the principle is *prima facie* or all things considered. So in nonassertoric contexts we will need an intrinsic indication of the status of the principle. Thus, indicators of closure seem to remain important logically, not just psychologically. And so we need to be willing to countenance negations of physical properties in our principles.

But by doing this, by introducing some elements that are not physical properties (namely, negations of physical properties), we are not introducing anything problematic. It is not as though we are introduc-

ing divine commands or queer metaphysical qualities into our description. The conditions are still limited to the physical and (occasionally) its absence.

By this provision, numbers will still not have physical descriptions because they do not have *any* physical properties (though they have a bunch of quysical properties). But maximal descriptions of people or actions with respect to their physical properties *will* be physical descriptions. The principles that result from supervenience will not introduce transphysical or only quysical properties. Post's metaphysical point would certainly protect us from this, too, but in far too sweeping a manner.

If we are interested in doing metaphysics, it is relevant and important to consider *what* we are talking about and why we are talking about it. To call "not being an electron" a physical description has much different implications when discussing numbers, which, after all, have no physical properties, from the implications of calling "not being bad tempered" a part of a physical (or, in this case, naturalistic) description when discussing Socrates, who, after all, has lots of physical (or natural) properties. It is condition (c) that ensures we are talking about something that already has physical properties.

It is a little-noticed fact about the supervenience condition we have assumed that it is consistent with the possibility of a wholly nonphysical entity having mental states, or moral qualities.<sup>10</sup> The supervenience condition simply assures that all nonphysical entities will have the same mental or moral characterization. (Since two nonphysical things have no physical differences between them, they can have no mental or moral differences between them.) And further, their mental or moral state is eternally unchanging.

As a matter of model theory (where we assume closure of the physical under complementation), this possibility will not block the entailment from supervenience to principles, because the condition of having no physical properties will count as a physical property, and thus be fit to serve in the antecedent of the conditional. But from Post's point of view, it will block the implication, since nonphysical things are, after all, nonphysical, and so can have no physical properties that will serve in an antecedent.

Again, however, in doing metaphysics, we need to attend to why we are doing it. Anyone who accepts the supervenience condition as it is stated believes that mental or moral qualities must be physically or naturalistically realized. Not only do the mental and the moral not vary independently of the physical, they cannot be instantiated in the absence of the physical. (I am not asserting that all philosophers of mind are physicalists. I am simply claiming that all philosophers of

mind who accept the supervenience of the mental on the physical are physicalists. Cartesians would have no reason to accept such supervenience in the first place.) So anyone who asserts the supervenience of the mental on the physical, for example, is also committed to the physical realization of the mental, and from these conditions follows the existence of physiocopsychical principles.

So when we ask ourselves the question that is at issue in my first problem, we have to decide whether we are asking it as model theorists or as metaphysicians. It is to Post's credit that he saw the need to ask the question as a metaphysician. But Post's metaphysical approach remains too abstract. For metaphysicians of mind or morality, certain presuppositions might be natural that could not be justified in a generalized metaphysical setting. Whether the set of physical properties is closed under complementation is a metaphysical question that depends on our intuitions and decisions about the nature of the physical (or the natural), as well as the context in which the question is being asked. Whether mental or moral qualities can be instantiated by entities that are nonphysical is also a metaphysical question. A requirement of physical (or naturalistic) realization would seem to be a presupposition of anyone who was willing to assert supervenience.

Since the idea of supervenience is most naturally used in the realm of ethics and philosophy of mind, it is reasonable to consult our thoughts about what function the notion of the physical is playing in those realms. And if we do so, I think it is natural to see the existence of principles as following from the assertion of supervenience. While metaphysical considerations seem to overturn conclusions reached from a model-theoretic point of view, deeper reflection on those metaphysical considerations takes us back to the original conclusions, but for much better reasons.<sup>11</sup>

## 2. The Second Problem

Some philosophers have been attracted by a form of supervenience known as global supervenience that seems to be weaker (in the sense of being less ambitious) than other forms of supervenience commonly considered. Global supervenience makes supervening properties depend not on the subvening properties of particular individuals, but on the distribution of subvening properties over the entire possible world. Thus, it is formulated as follows:

- (4) Possible worlds that coincide in respect of truths involving subvenient properties coincide in respect of truths involving supervenient properties.



And this is thought to be weaker than strong supervenience, which seems to localize the supervenience relation to the properties of particular individuals, as follows:

- (5) For any objects  $x$  and  $y$  and any possible worlds  $w$  and  $v$ , if  $x$  in  $w$  coincides in respect of truths involving subvenient properties with  $y$  in  $v$ , then  $x$  in  $w$  coincides in respect of truths involving supervening properties with  $y$  in  $v$ .

It is agreed all around that (5) implies (4). Kim originally claimed that (4) also implies (5) but more recently has conceded that it does not.<sup>12</sup> The following situation illustrates the failure of the latter implication.<sup>13</sup>

**World  $w$ :** There are exactly two objects  $x$  and  $y$  such that  $Px$ ,  $Mx$ ,  $Py$ ,  $\sim My$  (where  $P$  is a subvening property, and  $M$  is a supervening one).

**World  $v$ :** There are exactly two objects  $x$  and  $y$  such that  $Px$ ,  $\sim Mx$ ,  $\sim Py$ ,  $\sim My$ .

Possible worlds  $w$  and  $v$ , so described, constitute a model in which (4) holds but (5) fails. Thus (4) does not imply (5).

It is worth pondering what we mean by a "model." In particular, are there any constraints on the stipulation of models? After all, the situation described in  $w$  and  $v$  is rather odd. For if  $w$  and  $v$  are indeed possible worlds, it would seem that the following must also be possible worlds:

**World  $w^*$ :** There is exactly one object  $x$  such that  $Px$ ,  $Mx$ .

**World  $v^*$ :** There is exactly one object  $x$  such that  $Px$ ,  $\sim Mx$ .

So can we consider  $w$  and  $v$  to be a model in isolation from  $w^*$  and  $v^*$ ? If we insist that any model involving  $w$  and  $v$  must also include  $w^*$  and  $v^*$ , then the resulting more robust model fails to satisfy global supervenience after all.<sup>14</sup>

If we are only doing model theory here, then we can dismiss this proposed solution immediately: one needn't rule out the worlds  $w^*$  and  $v^*$ ; one simply doesn't include them in the stipulation. But perhaps there can be metaphysical constraints on the stipulation of models.<sup>15</sup> The problem with  $w$  and  $v$  seems to be that they try to paste inconsistent things together. The inconsistency is hidden by placing it in a larger context, but we can draw it out by a method of isolation. For example, consider the following restriction principle: any restriction of a possible world is itself a possible world.<sup>16</sup> One question would be, Under what conditions does it hold? A further question would be, Even if it is true, should it constrain the stipulation of models? If we allow this principle to constrain stipulation, then we force global super-

venience to be "fine-grained" in just the sense that its advocates wished to avoid.

Paull and Sider have proposed a way of reestablishing the independence of global supervenience as follows:<sup>17</sup> we simply need to construct a supervening property that cannot be subjected to the restriction principle for generating further possible worlds. They suggest the property  $M$ , which they define as holding of an object just in case it has  $P$  and some object has  $Q$ . (Intuitively,  $M$  is a nonintrinsic property.) Then  $M$  will supervene globally, but it will not supervene intrinsically on its possessor. The following situation is now supposed to illustrate the failure of (4) to imply (5):

**World  $s$ :** There are exactly two objects  $a$  and  $b$  such that  $Pa$ ,  $Qb$ ,  $Ma$ ,  $\sim Mb$ .

**World  $t$ :** There is exactly one object  $c$  such that  $Pc$  and  $\sim Mc$ .

And now there is no further possible world:

**World  $s^*$ :** There is exactly one object  $a$  such that  $Pa$  and  $Ma$ ,

because we know from the definition of  $M$  that (given only  $Pa$ )  $\sim Ma$ .

But it seems that if Petrie was guilty of ignoring further possible worlds that were entailed by his stipulated worlds, Paull and Sider are guilty of ignoring further subvening properties that are entailed by their stipulated properties. In particular, it seems that if  $P$  is a property, then there is also a property  $P\#$  that holds of an object just in case  $P$  holds of the object *and* there is another object in the world (formally:  $P\#x \equiv [Px \ \& \ (\exists y)y \neq x]$ ). (Intuitively,  $P\#$  is a nonintrinsic subvening property – i.e., it is a relational property.) So the correct description of the worlds would be:

**World  $s$ :** There are exactly two objects  $a$  and  $b$  such that  $Pa$ ,  $P\#a$ ,  $Qb$ ,  $Ma$ ,  $\sim Mb$ .

**World  $t$ :** There is exactly one object  $c$  such that  $Pc$ ,  $\sim P\#c$ ,  $\sim Mc$ .

So now, while this model does not violate (4), it does not violate (5) either, and so does not establish their independence.

Just as Paull and Sider urge us to transcend the purely formal approach to independence that Petrie takes and to consider the metaphysical question of which possible worlds there really are, we must also transcend their stipulative approach to the description of a possible world and consider the metaphysical question of which properties there really are (in the world). In doing so, we are brought back to the original model-theoretic conclusion, but for good metaphysical reasons.

A more radical challenge to the equivalence of (4) and (5) can be

gleaned from an article by John Haugeland.<sup>18</sup> He points out that what he calls "weak supervenience," essentially equivalent to (4), is consistent with the possibility that the truths in the supervening realm are about individuals that, as a set, are disjoint from the set of individuals that the subvening truths are about. Yet (5) requires that the supervening truths be about the same (domain of) individuals as the subvening truths. As long as the supervening domain is not constructible from the subvening domain, (4) can be true while (5) is not.

As a point about model theory, Haugeland is right. And he goes on to illustrate his point with examples drawn from plane geometry and wave physics. But a further question is what this means for the realms of mentality and morality.

Haugeland, in fact, argues that his point holds for mentality as well.<sup>19</sup> He uses, as a model, a chess-playing computer to which we ascribe intentional states. Haugeland claims that since the intentional states are a function of a large number of internal data structures, it is impossible to specify any particular data structures as *the* ones on which a certain intentional state (token) supervenes (or, with which it is token-identical).<sup>20</sup>

A response to this, however, can be constructed from the progress made on the first problem in this essay. An advocate of (5) can forsake the sort of specificity that Haugeland seems to require in a subvening characterization, and instead fall back on the idea of a possibly exhaustive characterization at the subvening level.<sup>21</sup> The key issue then turns out to be whether the individuals (or entities) to which mental or moral properties are ascribed can also be the subject of whatever properties are in question at the subvening level. It seems that they can, since the subvening level can just be treated as a different level of description of *that individual*. If that were *not* ultimately possible, then there would be a fundamental lacuna in our mental or moral practices.<sup>22</sup> The connection asserted between the levels by supervenience would seem to be nonlocalized and hence ungrounded. The possibility of a common domain assures this localizable grounding for the determination of one level by another.<sup>23</sup> We at least must be able to say that there is something (subvening and possibly relational) about these individuals that makes them whatever (supervenient) they are.

In assessing Haugeland's attempt to separate (4) and (5), we have had to focus on the nature of token-identity and on the notion of domains of individuals. Here our motivations have been mixed: I have objected to his narrow construal of token-identity, one that I think is too motivated by epistemological considerations, and replaced it with a wider construal. I am not sure if this takes us to metaphysics, or all the way back to model theory. But I have objected to his model-theoretic separation of domains on metaphysical grounds.

Research on supervenience over the past fifteen years has flourished both because of its theoretical elegance and because of its apparent metaphysical importance. It is time to articulate our motivations and recognize that they do not always point us in the same direction. Only then can we hope for progress.

## NOTES

In writing this essay I benefited from the stimulating ideas of Jaegwon Kim and the support of the NEH during Kim's NEH summer seminar entitled "Supervenience and Its Philosophical Applications" at Brown University, 1990. In revising the essay, I was helped by comments from Tom Grimes, John Heil, Bradford Petrie, Tadeusz Szubka, and Nick Zangwill. An unrevised version of part of the essay was presented to the 1992 meeting of the Southern Society for Philosophy and Psychology, Memphis, Tennessee, and was on that occasion subjected to helpful commentary by John Post.

1. What I am calling the "Quinean slogan" is approximated by Quine himself at several points in his essay "Facts of the Matter," *Essays on the Philosophy of W. V. Quine*, ed. R. Shahan and C. Swoyer (Norman: University of Oklahoma Press, 1978), pp. 162–6. The quotation from Donald Davidson appears in "Mental Events," *Essays on Actions and Events* (Oxford University Press, 1980), p. 214.
2. An assertion of *reducibility* entails that such conditions exist *and* are *necessary* as well as *sufficient*. I am not concerned with that more ambitious assertion. Of course, if there are not even sufficient conditions, reducibility will fail *a fortiori*.
3. R. M. Hare, "Supervenience," *Aristotelian Society Supplementary Volume*, 58 (1984): 1–16, at 3. See also p. 16.
4. Jaegwon Kim, "Concepts of Supervenience," *Philosophy and Phenomenological Research*, 45 (1984): 153–76, at 163–4. Following Kim, I have presented the issue in the particular case of whether weak supervenience implies principles. In fact, however, all the same considerations apply to strong supervenience as well. John Post has objected to formula (3), and by implication anything that (allegedly) implies it, as being excessively individualistic. Instead, he argues, we should endorse only global supervenience rather than (1) and (2). One might be able to surmount the apparent individualism of (1), (2), and (3) by allowing *P* to range over relational properties as well (see, e.g., Section 2). But whether that succeeds or not, the issue of interest to me – whether supervenience implies principles – arises for global supervenience too: from the assumption of global supervenience (no difference between two possible worlds without a physical difference between those worlds), does it follow that there is a physical description of (part of) a world that is a sufficient condition for (part of) the world to possess a supervenient property? Again, the same considerations apply. So, as I see it, the issue of individualistic versus global construals of supervenience (or, narrow

versus wide construals of mental states) is not relevant to the issue of whether supervenience implies principles.

5. John Post, *The Faces of Existence: An Essay in Nonreductive Metaphysics* (Ithaca, N.Y.: Cornell University Press, 1987), pp. 178–80.
6. Though one might have qualms about whether we can make sense of such an exhaustive characterization, which would involve considering a totality of physical properties, Post expresses no qualms about that. Insofar as not all subvening properties are “relevant” to the possession of certain supervening properties, it should be possible to get by with something less than an exhaustive characterization. Whether such a limitation could quell the qualms is a further question.
7. Post, *Faces*, p. 178. Post first made this point in his “Comments on Teller,” *Southern Journal of Philosophy*, 22, Supplement: *Spindel Conference on Supervenience* (1983): 163–7, at 165–6.
8. If, in the relevant contexts, the absence of a physical (or, natural) property could always be “represented” by the presence of some (other) physical property, then we wouldn’t have to involve ourselves with the metaphysical status of negations of physical properties at all. (In such contexts, negation would amount to choice-negation rather than exclusion-negation, since we would be dealing with objects that already had some physical properties.) But if, as seems more likely, precluding some physical properties cannot always be represented as (or, replaced by) including others, then we do need to consider the strategy proposed in the text.
9. While my essay supposes that closure will be achieved in the first way, Bradford Petrie has informed me that he prefers the second method, which he finds less ad hoc. Given a principle of the form  $(\forall x)(Px \rightarrow Mx)$ , where  $P$  is the conjunction of the subvening properties true of  $x$ , we would conjoin with it the following closure condition:
 
$$\forall y \{[(\forall Q) [(Q \in P) \rightarrow Qy] \ \& \ \neg(\exists R) [(R \in N) \ \& \ (R \notin P) \ \& \ Ry]] \rightarrow My\}$$
 (where  $Q$  and  $R$  range over subvening properties,  $M$  is the supervening property attributed by the principle,  $N$  is the set of all subvening properties, and  $P$  is the subset of  $N$  containing all subvening properties true of  $x$  – i.e., the set [not conjunction] of subvening properties attributed by the principle). I have pursued the first way because it seems to me that quysical properties do not deserve to be dismissed out of hand, but I do not reject the second. The second method seems purely model-theoretic.
10. But see, e.g., T. Grimes, “Supervenience, Determination, and Dependency,” *Philosophical Studies*, 62 (1991): 31–92, at 82.
11. The argument of this section is my attempt to substantiate a claim I made in note 4 of “Rationalism, Supervenience, and Moral Epistemology,” *Southern Journal of Philosophy*, 29, Supplement: *Spindel Conference on Moral Epistemology* (1990): 25–8, at 28. This section also corrects a mistake in the last paragraph of note 4 of my paper “Davidson’s Troubles with Supervenience,” *Synthese*, 85 (1990): 339–52, at 349.
12. Kim argued for the equivalence in “Concepts of Supervenience,” p. 168, and then conceded the inequivalence in “‘Strong’ and ‘Global’ Supervenience Revisited,” *Philosophy and Phenomenological Research*, 48 (1987): 315–26, at 318.

13. The example is taken from Bradford Petrie, “Global Supervenience and Reduction,” *Philosophy and Phenomenological Research*, 48 (1987): 119–30, at 121. Its use in this context is endorsed by Post.
  14. This way of reestablishing that (4) implies (5) was described by Jaegwon Kim in the seminar mentioned in the acknowledgment note above, though he does not necessarily endorse it. A similar strategy is employed by R. Cranston Paull and Theodore R. Sider in “In Defense of Global Supervenience,” *Philosophy and Phenomenological Research*, 52 (1992): 833–54, at 836.
  15. Kim writes, “David Lewis and Joseph Mendola have raised the possibility of using metaphysical considerations to disarm Petrie-type examples” (“‘Strong’ and ‘Global,’” p. 319n). Perhaps it was this sort of strategy they had in mind.
  16. See principle (1) in Paull and Sider, “In Defense,” p. 838. This principle strikes me as being analogous to insisting that  $S_5$  is the proper interpretation of modal logic. If one were to opt for a weaker relation of accessibility between worlds, then one could resist this principle by appealing to the limits of accessibility. An advocate of this principle seems to be committed to holding that the proper interpretation of modal logic is a metaphysical issue. If we follow this line of thought, we would probably wish to hold that the proper interpretation is relative to the domain about which modal assertions are being made. Then our metaphysical understanding of mentality, or morality, would be brought to bear in reflecting on the notions of mental, or moral, possibility.
  17. Paull and Sider, “In Defense,” p. 840.
  18. John Haugeland, “Weak Supervenience,” *American Philosophical Quarterly* 19 (1982): 93–103.
  19. Haugeland, “Weak Supervenience,” section 6. His point is specifically about token-identity, but the common-domain issue for supervenience is parallel.
  20. Haugeland writes, “. . . there is less reason to suppose that the metabolic constituents of mental events can be identified in the midst of all the irrelevant physiological housekeeping; and there is more reason to suppose that distinct mental events will each ‘supervene on’ (the activity in) extended brain regions, which may largely, or even entirely coincide” (“Weak Supervenience,” p. 101).
  21. The principles asserted to exist in Section 1 can be used to generate token-identities as follows: the instantiation of subvening properties in the antecedent of a principle are (token-)identified with the instantiation of a supervening property in the consequent. Haugeland’s notion of token-identity assumes a high degree of specificity at the subvening level, which, he is probably correct in assuming, could not be achieved. But the kind of token-identity that (5) commits us to turns out to be quite unspecific. Thus, I believe, Haugeland’s argument ultimately equivocates on the notion of token-identity: we should not accept a form of supervenience that commits us to token-identity (in the specific sense), but (5) commits us to token-identity (in the unspecific sense); therefore, we should reject (5).
- Once we understand the notion of unspecific principles and token-identities, we realize how little epistemological work they can do. (See my

essay "Rationalism, Supervenience, and Moral Epistemology.") Then the question beyond that becomes, How specific can they be made, and will they then be able to do any epistemological work? That is, I believe, an open question that cannot be answered by any model theoretic or metaphysical considerations. (Perhaps what I have shown in this essay is how far we can go if we are willing to lower our expectations enough.)

22. The scope of this argument is actually limited to supervenience claims involving those levels of subvening descriptions that are actually used in our moral or mental practices. A similar argument, but without this limitation, is given by Jaegwon Kim in "Supervenience for Multiple Domains," *Philosophical Topics*, 16 (1988): 129-50, sections 3 and 4. Roughly, Kim argues that, for a variety of reasons, we are committed to there being some kind of coordinating relation between the domains such that (4) implies (5).
23. Of course, the possibility of common (or coordinated) domains does not yet explain supervenience in a comprehensive fashion. Some would hold that where a supervenience claim does not reach as far as proper reduction, it is itself no more than a lacuna in our understanding. Still, some lacunae may be more tolerable than others. For more on these questions, see my essay "Wittgenstein and Neuroscience," *Synthese*, 78 (1989): 319-43, sections 6-10.