

# Does Same-Level Causation Entail Downward Causation?

Neil Campbell

Philosophy Dept., Wilfrid Laurier University,  
Waterloo, ON, N2L 3C5  
necampbe@wlu.ca

## Abstract

I argue that Jaegwon Kim's supervenience argument does not generalize to all special science properties by undermining his central intuition, employed in stage one of the argument, that there is a tension between horizontal causation and vertical determination. First, I challenge Kim's treatment of the examples he employs to support this intuition, then I appeal to Kim's own work on the metaphysics of explanation in order to dissipate the alleged tension.

A number of philosophers have discussed the concern that Jaegwon Kim's supervenience argument (1997; 1998; 2005) generalizes, thereby threatening to undermine the causal efficacy of *all* special science properties. For those who are attracted to a supervenience-based layered model of the world and who take seriously the idea that there are properties performing causal work on different levels, Kim's argument is clearly a threat. After all, Kim claims that the first stage of the supervenience argument shows that "*level-bound causal autonomy is inconsistent with supervenience or dependence between ... levels*" (Kim, 2005: 40). Discussions about whether or not the supervenience argument generalizes to other special science properties have focused primarily two points. Kim's critics have tried to show either that the fact the argument generalizes shows the argument is absurd because this implies causal drainage (Block, 2003), or else that Kim's suggestions for how to prevent causal drainage (Kim, 1997; 2003; 2005) are unsuccessful (Noordhof, 1999; Gillett and Rives, 2001; Bontly, 2002). Few, however, have critically examined the first stage of the supervenience argument itself. The reason for this, I suspect, is that the supervenience argument targets nonreductive physicalism, and most nonreductive physicalists are content to grant Kim his stage one conclusion that same-level mental-to-mental causation requires downward mental-to-physical causation. However, for those who are worried about the broader implications of the supervenience argument this failure to assess the first stage of the argument is an important oversight. As we shall see, there are grounds to reject the first stage of Kim's argument in which case the proponent of the layered model can avoid the conclusion that same-level causation *entails* downward causation. This means the supervenience argument might not generalize after all and that advocates of the layered model need not worry about causal drainage.<sup>1</sup>

---

<sup>1</sup>Of course, this strategy is unavailable to the nonreductive physicalist, and so my argument will not constitute a defense of nonreductive physicalism. This might make my discussion seem like a side issue but to characterize it as such is a mistake. For those who endorse the layered model and who conceive of levels in terms of mereological supervenience it is surely very important to preserve level-bound causal autonomy.

My discussion is divided into four parts. In part 1, I describe the tension Kim claims is at the heart of the first stage of the supervenience argument in the form of Edwards' dictum. In part 2, I outline the first stage of the argument and focus on the role that Edwards' dictum plays within it. In part 3, I attempt to dispel the alleged tension exploited in the first stage of the argument. My strategy is twofold. First, I identify some troubling features of Edwards' original example and an important disanalogy between what Edwards describes and the metaphysical framework of the supervenience argument. These observations raise some initial doubts about the reality of the tension Kim identifies in the first stage of his argument.<sup>2</sup> Second, I argue that the metaphysical and explanatory principles that might be used to motivate the tension actually have the opposite effect: they either undermine the tension, or else point to a way of relieving the tension that does not require downward causation. Either way, this will cast considerable doubt on Kim's claim that same-level causation *entails* downward causation and hence, on his broader conclusion about level-bound causal autonomy as well. In part 4 I consider, and ultimately reject, two ways in which Kim might respond to the above argument.

## 1 Edwards' dictum

The first stage of the supervenience argument involves the claim that there is a tension between same-level or "horizontal" causation and supervenience or "vertical determination." Kim introduces the tension by drawing on an example described by Jonathan Edwards, an eighteenth century philosopher and theologian. According to Edwards, God recreates the world at each instant *ex nihilo*. This means, contrary to appearances, that there are no temporally persisting objects or causal relations between such objects.

It will follow from what has been observed, that God's upholding created substance, or causing its existence in each successive moment, is altogether equivalent to an *immediate production out of nothing*, at each moment. Because its existence at this moment is not merely in part from God, but wholly from him, and not in any part or degree, from its *antecedent existence*. For the supposing that his antecedent existence *concurs* with God in *efficiency*, to produce some part of the effect, is attended with all the very same absurdities, which have been shewn to attend the supposition of its producing it *wholly*. Therefore the antecedent existence is nothing, as to any proper influence or assistance in the affair; and consequently *God* produces the effect as much from *nothing*, as if there had been nothing *before*. So that this effect differs not at all from the first creation, but only *circumstantially*. . . (Edwards, 1808: 331).

According to Edwards then, God's creative act, which can be thought of as a synchronic determinative relation, precludes the possibility of diachronic causal relations between objects. Kim adopts the general form of the tension that Edwards identifies, dubs it "Edwards' dictum," and defines it as follows:

*Edwards' dictum*: There is a tension between "vertical" determination and "horizontal" causation. In fact, vertical determination excludes horizontal causation (Kim, 2005: 36).

---

<sup>2</sup>Here my argument bears some similarity to one offered recently by Jens Harbecke (2013). However, Harbecke's discussion focuses on an anti-physicalist argument by Sturgeon (1998; 1999), involves a highly technical framework, and does not appear to work within Kim's own metaphysical assumptions. My hope is that my approach will be clearer and will be more compelling by virtue of working with assumptions and principles Kim himself endorses.

To illustrate the above tension and to clarify what he means by “vertical” and “horizontal,” Kim introduces a less esoteric example: the colour of a lump of bronze. If we want to understand why the bronze is yellow at a certain time,  $t$ , it seems we can appeal to its microstructure at that very time. That is, we can think of the colour as synchronically dependent on some subset of the microphysical properties of the bronze. Since the colour is a macroproperty in comparison to the bronze’s microstructure, we are to think of the micro-macro relation as one of “upward” determination, with the micro- and macroproperties arranged in a vertical array. The upward determination of the colour of the bronze is intended by Kim to be analogous to Edwards’ claim about God’s creative act, which is also a synchronic relation.

There appears to be another way one might account for the colour of the bronze, however: by appealing to its causal history. According to Kim, one might think of the bronze in terms of a series of successive “time-slices” and understand its colour at  $t$  as the causal product of the colour of the bronze at  $t-\Delta t$ , i. e., its colour at the previous time-slice. Since this involves a causal relation between objects and properties at the same mereological level, Kim calls this “horizontal causation,” with the causal arrow moving through time, from left to right. The diachronic causal relation between successive time-slices of the bronze should be thought of as analogous to the putative (and ultimately illusory) causal relation between temporally successive objects in Edwards’ example.

We appear to have a tension on our hands when we try to combine the two metaphysical accounts of the colour of the bronze at  $t$ . Kim sums the difficulty up as follows:

As long as the lump has microproperty  $M$  at  $t$ , it’s going to be yellow at  $t$ , no matter what happened before  $t$ . Moreover, unless the lump has  $M$ , or another appropriate microproperty (with the right reflectance characteristic), at  $t$ , it cannot be yellow at  $t$ . Anything that happened before  $t$  seems irrelevant to the lump’s being yellow at  $t$ ; its having  $M$  at  $t$  is fully sufficient in itself to make it yellow at  $t$  (Ibid.: 37).

The tension appears quite nicely to illustrate Edwards’ dictum. The upward determination of the bronze’s colour by its microstructure appears, as Kim says, to be in tension with the diachronic causation of its colour by an earlier time-slice of the bronze. In fact, so long as the bronze has the appropriate reflectance property at  $t$ , this being sufficient for its colour, the diachronic causal relation seems to be completely unnecessary, and so is pre-empted by the synchronic relation. Hence, with the aid of Kim’s own example, it would appear that Edwards’ dictum is a plausible one and identifies a genuine tension. Now that Edwards’ dictum is in reasonably clear focus, I will sketch out the first stage of Kim’s supervenience argument.

## 2 The first stage of the supervenience argument

Although our concern here is much broader, the intended target of the supervenience argument is nonreductive or “minimal” physicalism, including those forms of emergentism that are committed to the supervenience of emergent properties on their basal conditions.<sup>3</sup> The argument is a *reductio* that forces the nonreductive physicalist either to abandon her position (and embrace the reduction of mental properties to physical properties) or else to acquiesce in type epiphenomenalism. In Kim’s own words, the goal of the argument is to foist upon the nonreductive physicalist the following conditional thesis: “If mentality is to have a causal

<sup>3</sup>In fact, an earlier incarnation of the supervenience argument appears at the end of Kim’s paper, “Making Sense of Emergence” (1999). Since, in my view, the clearest and most thorough formulation of the supervenience argument appears in (Kim, 2005), my discussion will follow this version of the argument. This improves on the version in (Kim, 1998).

influence in the physical domain—in fact, if it is to have any causal efficacy at all—it must be physically reducible” (Ibid.: 161). Since I am only concerned with the first stage of the argument I will forego a lengthy treatment of Kim’s background assumptions and his alternative formulations of the second stage of the argument.

Stage one of the supervenience argument begins as follows: A putative causal relation between two mental events seems to require a causal relation between two mental property instances, such that

1. *M* causes *M\** (Ibid.: 39).

Since the nonreductive physicalist is committed to supervenience we have:

2. For some physical property instance *P\**, *M\** has *P\** as its supervenience base (Ibid.).<sup>4</sup>  
This is the point at which Kim appeals to Edwards’ dictum. He claims that (1) and (2) together give rise to a tension because we have two competing accounts of why *M\** is instantiated: *M\** is caused to instantiate by *M*, and *M\** is instantiated because its physical base *P\** is instantiated. Kim claims that the vertical determination of *M\** by *P\** appears to exclude the diachronic causation of *M\** by *M*. For what we see here is a tension analogous to the one between the synchronic determination of the colour of the bronze by its microstructure and the diachronic causation of its colour, or to the tension between God’s being the sustaining cause of the world and a diachronic causal relation between temporally successive objects. As with Kim’s other examples, if *P\** is instantiated, it doesn’t seem to matter what happened *before* the occurrence of *P\**. So long as the supervenience base *P\** is instantiated, *M\** *must* instantiate since *P\** is sufficient for *M\**.

According to Kim, in the light of Edwards’ dictum the only way to secure a causal role for *M* is to suppose that *M* caused *M\** to instantiate by causing its physical base *P\** to instantiate. The solution is elegant; since *M\** will instantiate as long as *P\** instantiates, the way to make *M* relevant to the instantiation of *M\** is to make it relevant to the instantiation of *P\**. This seems plausible since, to use Kim’s analogy (Ibid.: 20), in order to change the aesthetic qualities of a work of art one would need to alter its subvenient non-aesthetic properties; one cannot directly modify a painting’s beauty. Hence, we now have an appeal to what Kim calls “downward causation”:

3. *M* caused *M\** *by* causing its supervenience base *P\** (Ibid.: 40).

At this point in the argument Kim draws the following conclusions:

What the argument has shown at this point is that if *Supervenience* is assumed, mental-to-mental causation entails mental-to-physical causation—or, more generally, that “same-level” causation entails “downward” causation. Given *Supervenience*, it is not possible to have causation in the mental realm without causation that crosses into the physical realm. This result is of some significance; if we accept, as most do, some doctrine of macro-micro supervenience, we can no longer isolate causal relations within levels; any causal relation at level *L* (higher than the bottom level) entails a cross-level, *L* to *L* - 1, causal relation. In short, *level-bound causal autonomy is inconsistent with supervenience or dependence between the levels* (Ibid. 40).

---

<sup>4</sup>Since we are concerned here with the question of whether the supervenience argument generalizes we need only to think of *M* as a special science property and of *P* as its supervenience base.

I think the supervenience argument does not establish this conclusion about level-bound causal autonomy because I do not think Kim successfully motivates the move from (1) and (2) to (3) in the first stage of the argument. While it is true that most nonreductive physicalists are already committed to downward causation I don't think this commitment is or should be driven by the above argument. This is significant for the question of whether or not the supervenience argument generalizes—our main focus here. If I am correct and there are reasons to resist the above inference, then advocates of the layered model can preserve level-bound causation without the threat of those causal powers draining into their subvenient bases.

### 3 Revisiting the initial tension

To help bring my concerns into focus, I would first like to make some observations about Kim's treatment of Edwards's dictum. Let's begin with Kim's use of Edwards' example involving God as the sustaining cause of the world. It is important to think about the broader metaphysical picture and ask about the source of the tension between diachronic causation and God's synchronic creative acts. If we agree with Edwards that God recreates the entire world *ex nihilo* at every instant, this does appear to preclude the possibility of diachronic causal relations. But does the origin of the tension lie with the idea that it is problematic to have two sufficient metaphysical sources for how the world (or some part of it) came to be as it is at any particular moment? I think not. The real reason we are driven to reject diachronic causal relations in this case, I suspect, is that the metaphysical picture within which God's creative acts are embedded is one in which we lack a *necessary precondition* for diachronic causal relations. If we follow the tradition of thinking of *events* as the causal relata and embrace something like Lombard's (1986) view that events are changes in temporally persisting objects or substances, it is clear that the kind of world Edwards describes has no events since there are no temporally persisting objects. Hence, the real tension at work in Edwards' example might not originate from the idea that there are two sufficient metaphysical sources for the way the world is at any particular time-slice, but from the fact that there can be no events in the world Edwards describes, which means Edwards' dictum, as Kim articulates it, is perhaps somewhat misleading.<sup>5</sup>

In Kim's defense, Edwards' is a peculiar case, and it is tempting to overlook its rarefied metaphysical features and concede the point that something is amiss in the claim that it is possible to reconcile the idea that earlier events cause later ones if the latter are instantaneously brought into existence by God. Indeed, I am sympathetic with this suggestion. However, there is an important disanalogy between it and the tension Kim appeals to in the supervenience argument. In Edwards' example the competing metaphysical relations are both *causal relations*, so the source of the tension in this case (overlooking the above concerns about the overall metaphysical picture) is that we have two competing causes of the state of the world at any given time. If we take causes to be sufficient for their effects, as Kim does, this is certainly undesirable unless we are willing to endorse widespread causal overdetermination, which I assume we are not. It would be odd indeed to claim that God's omnipotent creative power is a redundant, overdetermining cause.

My reason for highlighting the fact that Edwards' example involves a tension between two competing causes is to point out that this is quite different from the alleged tension to which Kim appeals in the supervenience argument after premise (2). The upward determination of

<sup>5</sup>Such a world might, however, be consistent with the existence of Kimian events construed as property exemplifications, since temporal duration is arguably not an essential part of Kim's model (Kim, 1976). Of course, since Kim thinks (with Lombard) that events are the relata of causation, in a world without causation it might be difficult to motivate the existence of even Kimian events.

$M^*$  by  $P^*$  (unlike the relationship Edwards describes between God and the world) is *not* a causal relation, it is supervenience. So why should we agree with Kim that there is a tension between the claim that  $M^*$  is superveniently determined by  $P^*$  and the claim that  $M$  causes  $M^*$ ? If these were both causal relations there would indeed be a formal tension in need of reconciliation, and so we would have a reason to move to premise (3) of the supervenience argument. Since the situation is *not* one where we have two sufficient causes, why should we see *any* tension here?<sup>6</sup> It seems that the tension exists only if one conflates causal sufficiency with the sui generis sufficiency of supervenient determination. Since these are *different* metaphysical relations, why can't one claim they are compatible and simply deny the alleged tension?<sup>7</sup>

Kim claims in a footnote that he used to support (3) by appealing to an exclusion principle but that he now prefers "to rely on the reader's seeing the tension I spoke of in connection with the two answers to the question 'Why is  $M^*$  instantiated on this occasion?' ... I don't believe invoking any 'principle' will help persuade anyone who is not with me here" (Kim, 2005: 41). In the absence of any such principle, however, the question remains why one should see the tension, especially when it involves different kinds of metaphysical relations. Since the alleged tension is what drives the idea that same-level causation entails downward causation, surely it is not enough to justify it with an unsupported intuition. In the absence of any additional premises or background principles, the first stage of the supervenience argument threatens, at worst, to collapse at premise (2) or, at best, to lead to a stalemate between Kim and those who have a differing intuition about the compatibility of diachronic causation and supervenient determination. Thus, it is important to explore what additional metaphysical principles might break the stalemate and bring opposing intuitions in line with Kim's.

In what follows, I show that the additional principles at Kim's disposal fail to support the inference from (1) and (2) to (3) because they offer ways to alleviate the alleged tension that do not require an appeal to downward causation. If I am correct, Kim's argument does not justify the claim that "level-bound causal autonomy is inconsistent with supervenience or dependence between ... levels." In fact, Kim's other principles and general metaphysical framework suggest just the opposite.

In describing the alleged tension, Kim writes:

(1) and (2) together give rise to a tension when we consider the question "Why is  $M^*$  instantiated on this occasion? What is responsible for, and explains, the fact that  $M^*$  occurs on this occasion?" For there are two seemingly exclusionary answers: (a) "Because  $M$  caused  $M^*$  to instantiate on this occasion," and (b) "Because  $P^*$ , a supervenience base of  $M^*$ , is instantiated on this occasion" (Ibid.: 39).

In treating the two answers as "exclusionary" it looks like Kim is appealing to one of his infamous exclusion principles. Indeed, in stage two of the supervenience argument Kim appeals explicitly to *Exclusion*, which he formulates as follows:

*Exclusion*: No single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of causal overdetermination (Ibid.: 42).

Since *Exclusion* only generates a tension where there are multiple *causes*, yet we just observed that only one of the metaphysical sources of  $M^*$  involves a causal relation, the supervenience

<sup>6</sup>It is also interesting to note that Kim's use of Edwards' dictum is designed to *legitimize* diachronic causation (at the physical level)—something Edwards would clearly reject and regard as anathema to his own principle.

<sup>7</sup>Ausonio Marras (2007) also makes this observation but does not explore its implications since he is more concerned with the second stage of Kim's argument.

argument requires a broader principle than *Exclusion* in order to motivate the tension. Kim has two such principles at his disposal: the first is his well-known principle of *explanatory exclusion* (1988; 1989; 1995); the second is a still broader principle of *determinative/generative exclusion*, which Kim briefly mentions in his preamble to the supervenience argument (Kim, 2005). In the rest of this section I consider how these principles might operate in the first stage of the supervenience argument.

In a nearly identical version of the supervenience argument (1997), Kim appeals explicitly to the principle of explanatory exclusion in order to motivate the tension identified above. The principle is defined elsewhere<sup>8</sup> as follows:

*Explanatory exclusion*: there can be no more than a single *complete* and *independent* explanation of any one event, and we may not accept two (or more) explanations of a single event unless we know, or have reason to believe, that they are appropriately related—that is, related in such a way that one of the explanations is either not complete in itself or dependent on the other (Kim, 1988: 233).

If one were to accept such a principle, then there would indeed appear to be a tension involved in having two explanations for the occurrence of  $M^*$ . Furthermore, by having  $M$  cause  $P^*$ , which then determines  $M^*$ , as Kim proposes in premise (3) of his argument, one could alleviate the tension by showing that the two explanations of  $M^*$  are not independent of one another.<sup>9,10</sup> The causal dependence of  $P^*$  on  $M$  would *guarantee* that the explanations are not independent, so we can (at this stage of the supervenience argument) preserve them both by appealing to downward causation in the way Kim suggests. This is too quick, however. I think it important to consider Kim's treatment of the principle of explanatory exclusion and to look more carefully at how it might figure in the tension identified in the first stage of the supervenience argument.

I won't provide much in the way of detail about how Kim develops and defends the principle of explanatory exclusion, but a few points are central and deserve some elaboration. Kim motivates the principle by means of commitments to explanatory realism, causal realism, and some general considerations about the metaphysics of explanation. According to explanatory realism, a proposition  $C$  is an explanans for proposition  $E$  in virtue of there being a determinate, objective relation  $R$  between events  $c$  and  $e$ , that is, between the events the respective propositions are about.<sup>11</sup> Kim takes this relation to determine the correctness of an explanation and to serve as its objective content.

According to Kim, it is most often a causal relation that plays the role of the explanatory relation, hence the assumption of causal realism: that events stand in mind-independent causal relations with one another. Kim's support for the principle of explanatory exclusion and his intended meanings of the terms "complete" and "independent" that figure within it emerge from his survey of the various ways in which two causes ( $c_1$  and  $c_2$ ) of an event  $e$  might be

<sup>8</sup>Kim here refers to his (1989) in which he formulates the principle as follows: "two or more complete and independent explanations of the same event or phenomenon cannot coexist" but the principle is more perspicuous in his (1988).

<sup>9</sup>Graham MacDonald and Cynthia MacDonald (2006) exploit this fact and a particular version of the property exemplification view of events in order to undermine the supervenience argument.

<sup>10</sup>Note that we have here shifted from talking about the metaphysical *relations* involving  $M^*$  to talking about the *explanations* of  $M^*$ . In the present context this is harmless since, as we shall see, the independence of explanations is a matter of the independence of the explanatory relations they track.

<sup>11</sup>I will here follow Kim's convention of referring to events by lower case variables and propositions to the effect that those events occurred by the corresponding upper case variables.

related to one another. Suppose, in apparent contravention to the principle of explanatory exclusion, that we have two explanations for  $E$ : one in terms of  $C_1$  and another in terms of  $C_2$ . According to Kim, there are six ways in which the underlying causes might be related to one another: (1)  $c_1 = c_2$ , (2)  $c_1$  is distinct from  $c_2$  but is reducible to or supervenient on it, (3)  $c_1$  and  $c_2$  are partial causes of  $e$ , (4)  $c_1$  is a part of  $c_2$ , (5)  $c_1$  and  $c_2$  are different links in the same causal chain leading to  $e$ , and finally, (6)  $e$  is causally overdetermined by  $c_1$  and  $c_2$ . In case (1) Kim treats the explanations as equivalent. In cases (2) - (5), Kim claims that an explanation of  $E$  in terms of either  $C_1$  or  $C_2$  alone will be incomplete or fail to be independent of an explanation appealing to the alternative event. The reason an explanation of  $E$  in terms of  $C_1$  is *incomplete* under scenario (3), for example, is that  $c_1$  is on its own causally insufficient for  $e$ . Since  $c_1$  and  $c_2$  are each partial causes, any explanation that leaves out one of the causes will be incomplete. For Kim, given his commitment to explanatory realism, the metaphysical incompleteness of  $c_1$  in the production of  $e$  entails the incompleteness of the *explanation* of  $E$  in terms of  $C_1$ . Similarly, under scenario (4), an explanation of  $E$  that appeals to  $C_1$  fails to be *independent* of an explanation in terms of  $C_2$  because of the metaphysical dependence of  $c_2$  on  $c_1$ . Only in case (6), where  $e$  is causally overdetermined, are the alternative explanations complete and independent. Since causal overdetermination is sufficiently rare, Kim is happy to treat this as an exception to the exclusion principle.

Kim's catalogue of the various ways in which two explanations might be related provides plausible support for the principle of explanatory exclusion. Indeed, the principle of explanatory exclusion appears to be a straightforward corollary to the principle of causal exclusion appealed to in the supervenience argument. Given this, it would seem Kim is correct to maintain that we should treat the two explanations of  $M^*$  as not just *in tension*, but also as *exclusionary*.

This, however, is the point at which my earlier observations about Edwards' dictum become relevant once again. In the case of Edwards' example of God being the sustaining cause of the world, we can appreciate how a causal explanation for the way the world is at  $t$  that appeals to the previous time-slice is excluded by an explanation that appeals to God's creative act. Unless we are willing to endorse rampant causal overdetermination, only one cause can be sufficient, and hence, only one explanation can be maintained. I pointed out, however, that there is an important difference between Edwards' example and what is described in the supervenience argument. In the latter, the tension that Kim claims requires reconciliation via an appeal to downward causation involves two distinct kinds of relation: a diachronic causal relation between  $M$  and  $M^*$ , and a synchronic determinative relation involving the supervenience of  $M^*$  on  $P^*$ . Given this, it is not so clear that the principle of explanatory exclusion entails that these two explanations are exclusionary.

To sort this out we need first to recognize that Kim individuates explanations by the explanatory relation and the events it relates:

Explanatory realism yields a natural way of individuating explanations: explanations are individuated in terms of the events related by the explanatory relation (the causal relation, for explanations of events). For on realism it is the objective relationship between events that ultimately grounds explanations and constitutes their objective content (Ibid.: 233).

According to this criterion of individuation, we should see the  $M$ - $M^*$  relation and the  $P^*$ - $M^*$  relation as different explanatory relations that ground different explanations. Elsewhere, Kim (1994) argues that explanations track dependency relations, and he claims that there might

be many different kinds of dependency relations in the world. In fact, he treats mereological supervenience as an entirely distinct kind of explanatory relation from causation:

Another dependence relation, orthogonal to causal dependence and equally central to our scheme of things, is *mereological dependence* (or “mereological supervenience”, as it has been called): the properties of a whole, or the fact that a whole instantiates a certain property, may depend on the properties and relations had by its parts (Ibid.: 67).

Kim (1994) also claims that *mind-body* supervenience can be thought of as a different kind of dependency relation than causal dependence. Given Kim’s criterion of individuation for explanations, it is clear he should think the explanation of  $M^*$  that appeals to  $M$  is grounded in the diachronic causal relation between  $M$  and  $M^*$ , and tracks one kind of dependence, whereas the explanation of  $M^*$  that appeals to  $P^*$  is grounded in the supervenience relation, and hence, tracks a different kind of dependence. Assuming we adopt Kim’s principle of explanatory exclusion, how does it apply in this case? Does one explanation actually exclude the other? Since the latter is not a causal relation, as I suggested earlier, perhaps one can resist Kim’s argument by pointing out that there is no *formal* tension in need of resolution. In Kim’s own words, perhaps the explanation of  $M^*$  in terms of supervenience is simply “orthogonal” to the causal explanation of  $M^*$ . To the extent that the alleged tension is necessary to Kim’s conclusion that same-level causation entails downward causation, perhaps it is possible to block Kim’s move from (1) and (2) to (3) in the first stage of the supervenience argument. To settle this matter we need to delve a little deeper into the principle of explanatory exclusion and the potential justification for treating these two explanations as exclusionary.

The situation we face in stage one of the supervenience argument involves two different explanatory relations that converge on the same explanandum event ( $M^*$ ). In order to evaluate whether or not one excludes the other, we need to determine whether or not the explanations are complete or independent of one another. In Kim’s view, a causal explanation is complete when the cause it mentions is sufficient for the effect. In that case, the explanation that appeals to  $M$  in order to explain  $M^*$  would seem to constitute a complete explanation, given that  $M$  is assumed to be causally sufficient for  $M^*$ , but what of the explanation that appeals to the supervenient determination of  $M^*$  by  $P^*$ ? Is that also a complete explanation? In the light of Kim’s treatment of the example involving the colour of the bronze, it would seem so. Kim claimed that given the presence of the relevant base, the supervenient property *must* occur, no matter what happened at the previous time-slice. Hence, although a different kind of sufficiency is clearly involved in the case of supervenient determination, it seems the presence of  $P^*$  is (by strong supervenience) sufficient for  $M^*$ , and so the explanation of  $M^*$  in terms of  $P^*$  is also complete.

There appear, then, to be reasonable grounds to treat the two explanations as complete, but are they independent? It won’t do to claim the explanations fail to be independent simply because they converge on the same explanandum event. By that reasoning it would follow that explanations appealing to overdetermining causes aren’t independent—something Kim clearly denies. At first glance, it might seem unclear how we should understand the concept of independence that is relevant to the present example. After all, Kim’s discussion of the ways in which two explanations might fail to be independent appears to be cashed out primarily in terms of *causal* independence. However, two of his examples suggest a broader interpretation. Kim claims that causes fail to be independent if one supervenes on the other, or if one is a proper part of the other. It would seem, then, that causes fail to be independent if there is any

kind of metaphysical dependency relation between them—one that is more general than the relation of *causal* dependence that is active in cases of joint cause or of causal chains. This is hardly contentious as it amounts to a near tautology: given explanatory realism, explanations fail to be independent if the events they appeal to are not metaphysically independent of one another.

Are  $M$  and  $P^*$  independent in this sense? If we limit ourselves to the first stage of the supervenience argument there is simply not enough information to answer this question. To do so we need to consider the broader metaphysical picture that emerges in the second stage of the argument. If we grant that  $M$  has a supervenience base  $P$ , and that  $P$  causes  $P^*$ , then there are grounds for saying that the explanation of  $M^*$  that appeals to  $M$  is *not* independent of the explanation that appeals to  $P^*$  since  $M$  supervenes on  $P$  and  $P$  causes  $P^*$ . These dependency relations would appear to suggest that we could retain both explanations since the principle of explanatory exclusion allows for more than one explanation of an event, provided the explanations are appropriately related.<sup>12</sup> Since we are able to trace the explanatory relations that ground both explanations to a series of connected dependency relations, it appears that the two explanations *are* appropriately related, and so the principle of explanatory exclusion is not violated. Because the causal explanation of  $M^*$  can, in this way, be reconciled with the explanation of  $M^*$  that appeals to supervenience, it seems Kim's principle of explanatory exclusion does not support the claim that there is a tension in need of resolution in the first stage of the supervenience argument. We therefore lack any convincing reasons to suppose a tension exists, or if there is any *prima facie* tension, it is relieved simply by identifying the above dependency relations. Consequently, there is no need to appeal to downward causation in order to eliminate the alleged tension. We therefore have no reason to suppose that mental-to-mental causation entails mental-to-physical causation, or more generally, that same-level causation entails downward causation. We can, as Kim himself at one time suggested, regard the two explanations as “orthogonal” to one another rather than in competition.

If the principle of explanatory exclusion is unsuccessful at motivating the tension in the supervenience argument, perhaps Kim's other principle of determinative/generative exclusion will be more successful. In Chapter 1 of *Physicalism, or Something Near Enough*, Kim foreshadows the supervenience argument and proposes the following “generalized version of the exclusion principle”:

*Principle of determinative/generative exclusion:* If the occurrence of an event  $e$ , or an instantiation of a property  $P$ , is *determined/generated by an event*  $c$ —causally or otherwise—then  $e$ 's occurrence is not determined/generated by any event wholly distinct from or independent of  $c$ —unless this is a genuine case of overdetermination (Kim, 2005: 17).

Kim adds that the above principle expands on *Exclusion* because it “broadens causation, or causal determination, to generation/determination simpliciter, whether causal or of another kind” (Ibid.). By broadening exclusion to include *any* determinative relation, Kim would appear to succeed at reviving the tension in the supervenience argument between horizontal causation and vertical determination. With such a principle in place, perhaps one can no longer deny the tension in the first stage of the supervenience argument.

Kim's broader principle does make the alleged tension look more robust, but there are two ways to dissipate the tension that do not involve an appeal to downward causation: (1) reject

<sup>12</sup>Recall that the principle states: “we may not accept two (or more) explanations of a single event unless we know, or have reason to believe, that they are appropriately related—that is, related in such a way that one of the explanations is either not complete in itself or dependent on the other” (Kim, 1988: 233).

this broader version of exclusion; (2) argue as before that the two generative relations aren't independent of one another, and so we can preserve both.

The first strategy is obviously the strongest, but I will not belabour it here since what I have already said captures the essence of this reply. The main idea would be to show that, in the light of what we have already observed about Kim's views on the metaphysics of explanation, the broader version of the exclusion principle is unmotivated.<sup>13</sup> If one embraces explanatory realism and thinks of supervenience and causation as different *kinds* of explanatory relations, it is entirely unclear why we should embrace determinative/generative exclusion. The principle seems entirely ad hoc: Kim proposes it simply because it renders the tension in the supervenience argument more perspicuous. So the strategy here is to dig in one's heels and demand to know what is so objectionable about an event standing in two different kinds of explanatory relations. It would appear that, against the background of Kim's work on the metaphysics of explanation, the burden of proof lies squarely with him to show why we should accept this principle. It also bears pointing out that if, like Kim, one countenances causal overdetermination as a bona fide exception to determinative/generative exclusion, there is no *principled* obstacle to the convergence of distinct determinative or generative relations on a single event, and so it is fair to question the truth and value of such a principle, given that it has clear exceptions. The principle, in effect, seems to say "no overdetermination, unless this is a case of overdetermination" where "overdetermination" has been broadened to include any generative relation that is sufficient for the outcome. I fail to see the value of such a principle.<sup>14</sup>

Although I think one can deny the principle of determinative/generative exclusion, it is nevertheless possible to grant Kim this principle and still avoid the tension in a way that blocks the move from (1) and (2) to (3). We have already seen there are reasons for claiming the two metaphysical sources of  $M^*$  are not independent of one another once we consider the broader metaphysical picture within which  $M$ ,  $M^*$ , and  $P^*$  are embedded. The determinative/generative exclusion principle only threatens  $M$  as a cause for  $M^*$  if  $M$  and  $P^*$  are *independent*. Since, as we saw earlier, there are grounds for saying they are not independent, there is no reason to suppose that the determinative production of  $M^*$  by  $P^*$  excludes the causal generation of  $M^*$  by  $M$ . Nothing in Kim's principle suggests otherwise, and so the alleged tension evaporates once again.

## 4 Potential responses

I have argued that two versions of exclusion (explanatory exclusion and determinative/generative exclusion) fail to motivate a tension between vertical determination and horizontal causation, and hence, that there is as yet no convincing reason to think that same-level causation entails downward causation. Since we are working within Kim's own metaphysical framework it is worth asking if he has other resources that might provide a response to my argument. The seeds of two possible replies lie in Kim's earlier work on the metaphysics of explanation and on the concept of supervenience.

<sup>13</sup>It is interesting to note that Kim has his own doubts about this principle when he suggests that it might be too broad (Kim, 2005: 20).

<sup>14</sup>Of course, it might be implicit here that Kim thinks "genuine" cases of overdetermination are rare or that there are other reasons for thinking that we should not treat  $M^*$ 's being caused by  $M$  and being superveniently determined by  $P^*$  as a case of genuine overdetermination. While he has some plausible things to say about why we should not think that  $P^*$  is causally overdetermined by  $M$  and  $P$  (Kim, 2005: 46-52) it is not clear that those lessons will carry over to the set of relations now under consideration. I also think that Kim's claim that genuine overdetermination requires the overdetermining factors to be independent of one another is unmotivated.

The first involves an assumption about the individuation of explanations. Much of my defence of same-level causation requires the idea that it is possible for different explanatory relations to converge on a single event, and some philosophers might balk at such a suggestion. Indeed, Kim offers a brief remark on the subject that might support such a response. In footnote 23 of “Explanatory Realism, Causal Realism, and Explanatory Exclusion,” where Kim discusses how to individuate explanations, he writes:

If relations other than the causal relation can serve as [the] explanatory relation, they can also be considered as a basis for individuation; however, that probably would be redundant. It is unlikely that when the explanatory relation is different, exactly the same events would be involved (Kim, 1988: 239).

So perhaps one could argue that it would be highly unlikely that a single event can stand in more than one explanatory relation to other events. Since the denial of the tension in the supervenience argument requires something like the possibility Kim seems here to dismiss, it looks as though there might be a problem for the argumentative strategy I have proposed.

The trouble with this response is that the above remark seems to be entirely without support. Why must events stand in solitary explanatory relations with other events given that a variety of such relations are possible? Even Kim acknowledges that there are different kinds of explanatory relations in the world. More significantly, were there a principled reason to deny the possibility that one event can stand in two explanatory relations, that would rule out the possibility of causal overdetermination all by itself. Surely we are willing to allow the possibility of causal overdetermination, so why not also allow for the kind of case described above? To fail to do so seems arbitrary. Granted, in the case we are concerned with the explanatory relations are relations of different kinds, but that does not appear to be any barrier to their converging on the same event. In fact, one might claim that such a situation is *less* problematic than causal overdetermination precisely *because* the explanatory relations are relations of different kinds. Finally, it is worth noting that all Kim’s remark about individuation *technically* precludes is the *very same* pair of events standing in multiple explanatory relations. I agree with Kim that it would be extremely odd if, for example, *c* were not only a sufficient cause of *e*, but if *e* *also* strongly supervened on *c*. This, however, is not the case in the supervenience argument since the explanans for *M\** refer to two *different* events (*M* and *P\**). While *M\** stands in two different explanatory relations, each relation involves a distinct explanans-event.

A second possible response targets the claim, exploited above, that the two explanations of *M\** in terms of *M* and *P\** are not in tension because they are not independent of one another. Since *M* supervenes on *P* and *P* causes *P\**, *M* and *P\** are not independent of one another and so both explanations of *M\** can be preserved under explanatory and determinative/generative exclusion. Since we have stipulated that *M\** is determined by *P\** and that *P\** is caused by *P*, the only point where there might fail to be a dependency relation is between *M* and *P*. Why think there might not be a dependency relation here? The only justification I can think of involves the idea that supervenience itself fails to be a dependency relation. Though he was by no means the only one to raise this concern, Kim at one time (1995) argued that we lack an account of supervenience that is strong enough to secure dependence, yet weak enough to avoid reduction. If one shares Kim’s earlier concerns about the concept of supervenience, it might be possible to resist my argument by undermining the idea that the *M-P* supervenience relation is a dependence relation, thereby breaking the chain of metaphysical dependency between *M* and *P\**.

There is not much to recommend this strategy. This is so for two main reasons. First, such a claim would stop the first stage of the supervenience argument in its tracks anyway. This stage of the argument requires us to see a tension between the vertical determination of  $M^*$  by  $P^*$  and the causation of  $M^*$  by  $M$ . If we deny that supervenience is a dependency relation, we can no longer say that  $M^*$  is determined by  $P^*$ , and the tension evaporates. Hence, the argument would collapse at precisely the same point I am proposing, though for an entirely different reason. Furthermore, this would vindicate the same-level causal relation between  $M$  and  $M^*$  as the only genuine dependency relation in stage one of the argument, which is precisely the relation Kim wants to reveal as illusory.

I think this is a convincing reply, but another point against it deserves mention. That supervenience is a dependency relation is a key assumption of nonreductive or minimal physicalism—the main target of Kim’s argument. To see the relation as either reductive or as one of mere property covariation is, as Ruben (1990) points out, to fail to see it as an explanatory relation at all. This, of course, is simply to deny a key claim that defines the nonreductive physicalist’s position. If stage one of the supervenience argument requires such a denial as a *background assumption* in order to block my response, the argument is certainly question begging. For to assume a deflationary view of supervenience is already to deny the nonreductive physicalist’s position, since she takes this dependence to define her position and to legitimize her claim to be a physicalist. Hence, this does not appear to be a plausible way to respond to the argument.

In conclusion, if one does not see a tension between the diachronic causation of  $M^*$  by  $M$  and the supervenient determination of  $M^*$  by  $P^*$ , there is no need to appeal to downward causation in order to alleviate the tension. Only if the tension is genuine should we be compelled to think that mental-to-mental causation requires mental-to-physical causation, or that same-level causation entails downward causation more generally. I hope to have shown that the tension between same-level causation and supervenient determination is unproblematic, and have done so working from within Kim’s own metaphysics. If I am correct, advocates of the layered model need not worry that the supervenience argument will generalize or that the efficacy of special science properties will drain into their physical bases.<sup>15</sup>

## References

- Block, N. (2003). “Do Causal Powers Drain Away?”, *Philosophy and Phenomenological Research* 67, 133–50.
- Bontly, T. (2002). “The Supervenience Argument Generalizes”, *Philosophical Studies* 109, 75–96.
- Edwards, J. (1808). *The Doctrine of Original Sin Defended: Evidences of Its Truth Produced, and Arguments to the Contrary Answered. Containing in Particular, a Reply to the Objections and Arguings of Dr. John Taylor, in His Book Intituled, "the Scripture Doctrine of Original Sin Proposed to Free and Candid Examination"*, Boston: S. Kneeland.
- Gillett, C. and Rives, B. (2001). “Does the Argument from Realization Generalize? Responses to Kim”, *Southern Journal of Philosophy* 39, 79–98.
- Harbecke, J. (2013). “On the Distinction between Cause-Cause Exclusion and Cause-Supervenience Exclusion” *Philosophical Papers* 42, 209–38.
- Kim, J. (2003). “Blocking Causal Drainage and Other Maintenance Chores with Mental Causation” *Philosophy and Phenomenological Research* 67, 151–76.

<sup>15</sup>I wish to express my gratitude to the anonymous referees for this journal for their helpful comments on an earlier draft of this paper.

- Kim, J. (1997). "Does the Problem of Mental Causation Generalize?", *Proceedings of the Aristotelian Society* 97, 281–97.
- Kim, J. (1976). Events as Property Exemplifications, in Brand, M. and Walton, N., eds, 'Action Theory', Dordrecht: Reidel. Reprinted in Kim, J. 'Supervenience and Mind', Cambridge: Cambridge University Press, 1993). References are to the 1993 reprint.
- Kim, J. (1995). Explanatory Exclusion and the Problem of Mental Causation, in MacDonald, C. and MacDonald, G., eds, 'Philosophy of Psychology: Debates on Psychological Explanation', Oxford: Blackwell, 35–56.
- Kim, J. (1994). "Explanatory Knowledge and Metaphysical Dependence", *Philosophical Issues* 5, 51–69.
- Kim, J. (1988). "Explanatory Realism, Causal Realism, and Explanatory Exclusion" *Midwest Studies in Philosophy* 12, 225–40.
- Kim, J. (1999). "Making Sense of Emergence", *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 95, 3–36.
- Kim, J. (1989). "Mechanism, Purpose, and Explanatory Exclusion", *Philosophical Perspectives* 3, 77–108.
- Kim, J. (1998). *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*, Cambridge, Mass.: MIT Press.
- Kim, J. (2005). *Physicalism, or Something near Enough*, New Jersey: Princeton University Press.
- Lombard, L. B. (1986). *Events: A Metaphysical Study*, London; Boston: Routledge & Kegan Paul.
- Macdonald, C. and Macdonald, G. (2006). "The Metaphysics of Mental Causation", *Journal of Philosophy* 103, 539–76.
- Marras, A. (2007). "Kim's Supervenience Argument and Nonreductive Physicalism", *Erkenntnis* 66, 305–27.
- Noordhof, P. (1999). "Micro-Based Properties and the Supervenience Argument: A Response to Kim", *Proceedings of the Aristotelian Society* 99, 109–114.
- Ruben, D. (1990). *Explaining Explanation*, London: Routledge.
- Sturgeon, S. (1999). "Conceptual Gaps and Odd Possibilities", *Mind* 108, 377–80.
- Sturgeon, S. (1998). "Physicalism and Overdetermination", *Mind* 107, 411–32.