SUPERVENIENCE AND EXPLANATORY EXCLUSION∗

LEE McINTYRE
Harvard University
lee_mcintyre@harvard.edu

SUMMARY: This paper argues that there is an inconsistency between Jaegwon Kim’s earlier work on supervenience and his more recent work on explanatory exclusion. In his work on supervenience Kim advocates an explanatory agnosticism that, by the time of his later work, is replaced by an endorsement of reductive explanation. My argument is that this tension between Kim’s early and later work is unfortunate since explanatory exclusion is highly questionable in its own right and is not reconcilable with his earlier work on supervenience anyway.

KEY WORDS: supervenience, explanatory exclusion, reductionism, Jaegwon Kim

RESUMEN: El presente artículo sostiene que existe una inconsistencia entre los primeros textos sobre superveniencia de Jaegwon Kim y sus trabajos más recientes sobre exclusión explicativa. En sus escritos sobre superveniencia, Kim defiende un agnosticismo explicativo que, en sus textos posteriores, sustituye con su aprobación a la explicación reductiva. Mi argumento es que esta tensión entre las dos etapas de la obra de Kim es poco afortunada, pues la exclusión explicativa es por derecho propio sumamente cuestionable y de todos modos no se puede conciliar con sus primeros textos sobre superveniencia.

PALABRAS CLAVE: superveniencia, exclusión explicativa, reduccionismo, Jaegwon Kim

I

For over a decade Jaegwon Kim’s work on supervenience has stood as a model of clarity and analytical rigor for those interested in metaphysics and the philosophy of mind. In a series of recent articles on the doctrine of “explanatory exclusion” Kim has taken up a new set of questions surrounding the conditions of explanatory adequacy, in the hope of following up some of the consequences of his earlier work, and settling the epistemological issues as admirably as he has the metaphysical ones. It is my

∗ I would like to thank Terry Horgan, Jaegwon Kim, Ausonio Marras and an anonymous referee for this journal for their helpful comments on an earlier draft of this paper.
contention, however, that there is a deep inconsistency between Kim’s earlier work on supervenience and his more recent work on explanatory exclusion.

In reading Kim’s work on supervenience, hints of an explanatory program can already be discerned. For from the outset, the question that has haunted supervenience is whether or not it implies reductionism. In his first and most comprehensive account of the supervenience relationship, Kim makes the following statement:

strong supervenience [...] says nothing about how successful we shall be in identifying causes and framing causal explanations; it is also silent on how successful we shall be in discovering causal laws. Explanation is an epistemological affair [...]. [T]he thesis that a given domain supervenes on another is a metaphysical thesis about an objectively existent dependency relation between the two domains; it says nothing about whether or how details of the dependency relation will become known so as to enable us to formulate explanations [or] reductions.¹

In the same article, Kim further addresses the issue of whether supervenience implies reductionism when he argues that: “[Supervenience] acknowledges the primacy of the physical without committing us to the stronger claims of physical reductionism.”² And, in case there was any doubt, Kim addresses the issue a third time, stating that: “Reduction, explanation, and the like are epistemic activities, and the mere fact that […] equivalence or biconditionals ‘exist’ is no guarantee that they are, or will ever become, available for reductive or explanatory uses.”³

Thus, Kim’s position on explanation is here quite unambiguous. He argues that there is a break between metaphysics and epistemology; that it is one thing to say that a metaphysical relationship like supervenience may potentially support reductive explanation and quite another to say that it compels it.

¹ Jaegwon Kim 1984, p. 175.
² Kim 1984, pp. 155–156.
It is worth lingering for a moment here, in an attempt to set out Kim’s apparent position on explanation in more detail, for it is precisely on this basis that the later tension in Kim’s work will become manifest. What is Kim’s position on explanation—and in particular on reductive explanation—in his work on supervenience? It seems to be this: while Kim seems comfortable with the view that supervenience supports the in principle possibility of reductive explanation—given certain conditions that may govern the particular metaphysical relationships that we are concerned with and our knowledge about them—what he is not willing to support is the idea that reductive explanation is necessary when supervenience applies. As we have seen in the above quotations, Kim seems at pains to point out (1) that even strong supervenience does not imply that these conditions obtain, and (2) that even if they did we would not necessarily know enough about them to effect a reduction. Thus, in any meaningful sense, Kim’s early views on supervenience seem to put him strongly on the side of the debate which says that supervenience does not imply reductionism.

II

In considering Kim’s more recent work on explanatory exclusion one realizes that his views on explanation have evolved in a way that seems at odds with his earlier commitments. For in a series of articles: “Mechanism, Purpose, and Explanatory Exclusion” (1989a), “Explanatory Realism, Causal Realism, and Explanatory Exclusion” (1987), and “The Myth of Nonreductive Materialism” (1989b), the position on explanation that Kim develops is that where material dependency relationships (like supervenience) exist there is one and only one genuine explanation that can be given for any event, and this type of explanation is reductive.

In “Mechanism, Purpose, and Explanatory Exclusion”, Kim defines explanatory exclusion as the principle that “no event can be given more than one complete and independent expla-
nation”. 4 Immediately recognizing the ambiguity of these qualifiers, Kim goes on to write:

What “complete” and “independent” may mean in this context is obviously important, and my discussion will be sensitive to the need of making these notions clearer; I should say right now, though, that I shall not be offering general definitions of these notions, but depend rather on the discussion of specific cases to generate reasonably cohesive senses for these terms. 5

One such case, immediately following Kim’s proposal to employ this strategy, is a discussion of the relationship between “rationalizing” and “physiological” accounts in the explanation of psychological phenomena. 6 Here Kim claims that, in those cases where phenomena at the secondary level of description (constituting the “psychological event”) have a nomological correlate at the primary level (constituting the “physiological event”), “all the conditions necessary for the reduction of psychology to physiology [...] are satisfied”. 7 The metaphysical dependence between the two levels of inquiry (here read as nomological) makes the independence of the psychological level of explanation suspect and therefore a candidate that is “ripe for reductive absorption”. 8

Coming full circle to his initial condition of “independence” that must be met by an adequate explanation, Kim says that in this instance “if the rationalizing [secondary] explanation is dependent on the physiological [primary] one in an appropriate sense (i.e., by being reducible to it), then in truth there is only one explanation here [...].” 9 A few pages later, Kim concludes tersely that in such cases of material dependence, “the two ex-

---

4 Kim 1989a, p. 79.
5 Kim 1989a, p. 79.
6 Kim 1989a, p. 80.
7 Kim 1989a, p. 88.
8 Kim 1989a, p. 88.
9 Kim 1989a, p. 80.
planations are no longer independent—one is reducible to the other." 10

It is thus clear how, in his later work, Kim interprets material dependence as a proxy for reductive explanation. We here see how Kim’s views on “independence”, in his definition of explanatory exclusion, commit him to the idea that where material dependency exists there is one and only one (complete) explanation that can be given for any event, and that that explanation must be reductive. Material dependence, by this logic, inexorably leads to explanatory exclusion, for the only type of explanation that can be “independent” is one that captures the causal mechanism behind the phenomena in terms of their most basic material counterparts —i.e., a reductive explanation.

Of course, as in his earlier statements on supervenience, Kim may here wish to submit that we do not know enough right now to effect some reductions, even in those cases where we have strong evidence of ontological dependence. But, what now seems sanctioned is the idea that reductive explanation is an ideal type, that should be aimed at whenever possible. Indeed, to put the point in its strongest terms, Kim now appears to believe that we have not truly given an explanation at all unless we have reduced all secondary explanations to their most basic level of material expression. 11 Thus, even while Kim recognizes that there may be some situations in which no explanations can be given (because of our epistemological limitations), where an explanation can be given it must be reductive.

10 Kim 1989a, p. 88. Recall that, following Kim’s strategy, this example is intended to be emblematic of the conclusions that should be drawn about explanation for all such cases of material dependence.

11 Kim embraces just such a principle when he rejects “vernacular psychology” in favor of neuroscience in Kim 1989a, p. 101, where he writes: “The explanatory exclusion principle provides a simple explanation of why the two theories [. . . ] compete against each other and why their peaceful coexistence is an illusion. For vernacular psychology and neuroscience each claim to provide explanations for the same domain of phenomena [. . . ]. Hence, by the exclusion principle, one of them has to go.” In an accompanying footnote, Kim makes it explicit that the victor here is neuroscience, but offers that—if one still wishes to preserve vernacular psychology—perhaps it would be better not to think of its goal as explanation.
The incommensurability between Kim’s earlier and later views about explanation here can be seen quite sharply. Where- as Kim’s earlier views seemed to imply that supervenience need not compel reductive explanation, his more recent writings on explanatory exclusion would seem to commit him precisely to the claim that it must. By the time we get to the doctrine of explanatory exclusion, supervenience is disguised reductionism, or might just as well be.

In the face of these charges, perhaps one might try to save Kim by arguing —as he does himself— that supervenience implies only the possibility of reductionism; that it merely provides the metaphysical conditions that would make such explanations allowable.\(^{12}\) In pursuit of the above strategy, Kim has recently written that:

we can perhaps begin with a distinction between (a) the claim that the mental supervenes on the physical and (b) an actual supervenience scheme (which exhaustively specifies which mental properties supervene on which physical properties, for all mental properties). The supervenience claim (a) may very well be true, without our ever being able to produce (b), a supervenience scheme which is even remotely complete.\(^{13}\)

Yet, one questions the effectiveness of such a strategy in rescuing Kim’s views. For in the above statement Kim writes as if the charge against him were that supervenience implies that reductive explanation is forthcoming. Yet, if one examines more closely Kim’s work on explanatory exclusion, a much stronger (and more controversial) claim seems to be advocated: that where metaphysical relationships like supervenience obtain, only reductive explanations will be found to be satisfactory (whether they are forthcoming or not)!\(^{14}\) Indeed, isn’t the point of explanatory exclusion to highlight for us the extent to which ontological factors must constrain explanatory accounts, even if we cannot guarantee that such explanations ever will become

---

\(^{12}\) Kim, personal communication, August 4, 1994.

\(^{13}\) Kim, personal communication, 1994.

\(^{14}\) Kim 1989a, pp. 87–88.
available? Thus, does explanatory exclusion seek to define the parameters for what sort of explanations are acceptable where strong metaphysical relationships —such as supervenience— apply, and tells us that in such cases we ought to pursue reductive explanations. Even if one concedes the point that supervenience cannot show that a reductive explanation must be forthcoming, it is nonetheless troubling to think that where supervenience relationships obtain reduction is our only explanatory option.

I will not rehearse here the familiar arguments against reductive explanation. I will simply note that the view that only reductive explanations are truly explanatory —which seems at the heart of the doctrine of explanatory exclusion— is one that is quite narrow and prejudicial in its view of the parameters of explanatory adequacy and must be defended in a way more searching than to argue that relationships like supervenience make it theoretically possible.

What is of more immediate interest for our present analysis, however, is how such a prejudice follows (if it does) from any commitments about supervenience. Does supervenience in and of itself lead to explanatory exclusion? Are they even reconcilable? This question, perhaps, is only answerable if we now bring to the table what I argue is an additional explanatory constraint—itself highly questionable—that seems to underlie Kim’s principle of explanatory exclusion: the doctrine of “descriptivism”.

15 An excellent discussion of reductionism and its problems can be found in Lawrence Sklar 1967. The classic source on reductionism is Ernest Nagel 1961, chapter 11.

16 Kim has recently denied adherence to the doctrine I have defined as “descriptivism”, claiming that in his earlier work he has made clear that there is a distinction between “supervenience as a metaphysical relation and reduction as an explanatory activity”. Yet this is precisely my point in arguing that his earlier and later works are inconsistent, for while he does seem to reject the doctrine of descriptivism in his work on supervenience, in his work on explanatory exclusion he appears to implicitly accept at least some qualified version of it. (Personal communication, March 27, 1996.)
II

Elsewhere I have presented “descriptivism” as the basis for a dubious view about the impossibility of offering autonomous nomological explanations in those cases where relationships like supervenience obtain.\textsuperscript{17} “Descriptivism” can be defined as the view that the subject matter of any given science is not the phenomena presented to us by nature, but the phenomena as captured by some particular vocabulary.\textsuperscript{18} It is the view that one level of description of the phenomena is constitutive of the subject matter of a discipline. Such a view, I have argued, paints a narrow picture of the relationships that we think it is legitimate to desire to have explained and artificially limits us to one particular way of capturing them. Yet it is adherence to the doctrine of descriptivism, I claim, that plausibly lies behind Kim’s doctrine of explanatory exclusion.

Supervenience alone, or really any particular view about the nature of metaphysical dependency, could not by itself motivate the strong and restrictive set of parameters that characterize the doctrine of explanatory exclusion. Given the break between metaphysics and epistemology, Kim’s earlier agnosticism about the explanatory consequences of supervenience just seems right. One may rationally adhere to the doctrine of supervenience, that is, and yet reject the doctrine of explanatory exclusion, even though Kim does not. Why doesn’t Kim? I argue that it is because he tacitly accepts the doctrine of descriptivism, or some qualified version of it, which artificially limits the horizon of explanatory accounts that are deemed acceptable in cases of strong material dependence.\textsuperscript{19} For what Kim rejects in his ac-

\footnotesize
17 McIntyre 1996.
18 Cf. note 21 below.
19 That Kim implicitly accepts some additional principles to get him from supervenience to explanatory exclusion is made clear in Kim 1989a, p. 89, where he writes “A thorough examination of explanatory exclusion will inevitably spill over into the long-standing debate over the nature of explanation, a topic on which nothing like a consensus now exists. The discussion to follow [in which Kim hopes to show that explanatory exclusion is a reasonable constraint on all causal explanations] will inevitably rest on certain intuitive assumptions [emphasis added] about how explanations [. . .] work.”
count of explanatory exclusion seems identical to what is scorned by descriptivists: that one may offer multiple explanations of one and the same material relationship, since there are many different ways of describing any given metaphysical connection. That Kim rejects this idea is not in dispute; he says as much.\textsuperscript{20} The only remaining question to be considered, then, is whether descriptivism is a reasonable assumption and whether it plays a legitimate role in bridging the gap between supervenience and explanatory exclusion. We shall start with the second claim.

IV

The doctrine of explanatory exclusion seems a queer consequence to draw from supervenience. Now, it is true that one need not draw it directly in order to back up Kim’s claim that they are compatible. But the discontinuity with Kim’s earlier statements on the epistemological consequences of supervenience here comes back to haunt him, for it does seem that the appropriate explanatory conclusion to draw from supervenience alone is one of agnosticism. By the time of his statements about explanatory exclusion, however, Kim is not at all agnostic about the type of explanation that is appropriate in those cases where supervenience obtains; he tells us that we should seek reductive explanations.

It is worth noting here, irrespective of the connection to supervenience, just how radical a claim is made by the doctrine of explanatory exclusion. For what explanatory exclusion denies is

I believe that a plausible candidate for one of these intuitive assumptions is descriptivism.

\textsuperscript{20} In Kim 1989a, pp. 94–97, he nests his account of explanatory exclusion within a framework of “explanatory realism”, whereby he accepts the notion that there is one true and objective fact about what causes any specific event and that there is one and only one correct way to describe it, which forms the basis for its explanation. We here see how descriptivism and explanatory realism—which is another of his tacit assumptions—go hand in hand. Cf. here Kim 1987, p. 239, in which he admits that explanatory realism is a tacit assumption of explanatory exclusion. That explanatory realism ignores the question of what description of an event is taken to be “objectively true” only reinforces the case that Kim’s epistemology is descriptivist.
precisely something that we often find valuable in defining the conditions of explanatory adequacy—that despite the strength of the ontological connection between different levels of organization, we seek to pursue explanations at many different levels of description, because there are many different ways of describing even one and the same ontological relationship. Explanation is not of the phenomena as such, but as described by a certain theoretical vocabulary. And what seems odious about the desire to sanction only reductive explanations, or to claim that no explanation can be given until we do become capable of reducing relationships to their basic causal components, is that we run the risk of eclipsing the very explanandum that we sought to explain in the first place.

How might one attempt to reconcile the tension between Kim’s views on supervenience and explanatory exclusion? First, one might try to argue that while supervenience does not compel reductionism, reductive explanation is the only type that we would find adequate in those cases where supervenience obtains. That is, one might claim that while supervenience does not imply anything about the actual possibility that we will ever be able to offer reductive explanations, the governance of the explanandum by supervenience does indeed imply that a reductive explanation is the only one that will be acceptable. But this view seems odd absent a more overt connection to some doctrine that accounts for why supervenience should have such strong normative force in shaping our explanatory commitments. In short, why would the theoretical possibility of reductionism, which Kim earlier admits is afforded by supervenience, suddenly drive the claim that reductive explanation is the only kind acceptable where supervenience governs? It must be that alternative levels of description of an ontological relationship are not accepted as the basis for genuine explanatory accounts. Autonomous explanations are therefore rejected. The phenomena at the secondary

21 It is worth noting the distinction between this claim and the doctrine of “descriptivism”. To claim that events can be explained only when they are described in some way does not mean that the subject matter is exhausted by the description of the phenomena in a particular way.
level may be expressible in terms other than those used at the primary level on which they are dependent —this view asserts— but these may not form the basis for autonomous explanations. Yet this commitment just seems tantamount to descriptivism!

Now, perhaps one might here wonder why it couldn’t simply be held that while supervenience alone does not imply reductionism, supervenience along with other explanatory constraints (like descriptivism) does. But, if this strategy is followed, one wonders why in Kim’s earlier writings such a fuss was made over the fact that supervenience does not imply reductionism. If Kim believes —on other grounds— that where supervenience relationships obtain reductive explanations are the only ones that will be adequate, why not say so explicitly rather than leave the impression that he has changed his mind since his earlier writings on supervenience? Either (1) Kim now thinks that supervenience does imply reductionism, or (2) he never thought that it did, but instead believes on independent grounds that reductive explanations are the only type adequate where supervenience obtains. If the former view is correct, Kim owes it to us to say that he has changed his mind. But, if the latter, what precisely are these additional commitments that propel us toward explanatory exclusion, and why not identify them explicitly?

V

I hope already to have shown that supervenience by itself does not land us at the door of explanatory exclusion. This is because supervenience is not disguised reductionism. But has it therefore been shown that Kim’s views on supervenience and explanatory exclusion are inconsistent? Even though he does not identify it as an explanatory constraint, might not supervenience and explanatory exclusion be reconcilable once one accepts the doctrine of descriptivism? Is the notion of “descriptivism” defensible in its own right?

I believe that it is not. Descriptivism is an artificially narrow view about the conditions governing those explanations that we are willing to accept as adequate. It supports the desire to reject autonomous levels of description as a foundation for nomological
—or any other kind of— explanation at secondary levels. What is the basis for this doctrine? Simply the idea that we already have in hand (or have confidence that we will soon have in hand) the correct descriptive terms that will govern our explanation of the explanandum. But how do we know this? Indeed, isn’t this just the sort of thing that Kim, in his earlier writings on supervenience, said that we could never be sure about? As a practical matter, how can we be sure that we have ever obtained, or ever will obtain, the correct descriptive terms governing any secondary (or even primary) relationship?

Furthermore, over and above any practical problem that descriptivism may face, I submit that there is an even more fundamental embarrassment facing this account. For often Kim (and other descriptivists) speak as if the explanations they seek can be given in God’s ideal Adamic language, with no translation into a partial or limited descriptive account whatsoever. Yet how is it possible at the primary level any more than it was at the secondary level to effect such a perfect translation of the causal engine that drives nature? Don’t reductive explanations too rest upon descriptions? But by what translation? To which reductive base? By what set of bridge laws?

There seems to be a conceptual confusion at the heart of the desire to forego explanation at all secondary levels of description, in favor of the pursuit of reductive explanation, so that we can get our explanations to capture the most basic facts of ontological dependency. For even reductive explanations are dependent upon our descriptions of reality and are theory dependent. How then can we hope to get to the root of ontological dependency by using reductive explanations? Isn’t there more than one way of characterizing such relationships, and therefore many possible competing reductive explanations? Indeed, even if we do not believe so, and simply must wait until we have the ideal vocabulary for capturing ontological dependency, how may we ever hope to explain anything in the interim?

The goal of reducing all explanations to a single foundational account is a myth. For the descriptivist assumption is just as false when one is aiming at reductive explanations as when one
is offering autonomous ones. There is no single objective vocabulary that even in principle can be used to capture the true nature of ontological dependency. Our explanations of material relationships must always be mediated by an incomplete and imperfect language. Even if one is a committed materialist, who believes that only reductive explanations are adequate, we can never hope to escape the fact that there are many different ways of describing one and the same reality, each forming the basis for alternative explanatory accounts. Descriptivism, that is, is false precisely because, once we realize that our knowledge of ontology is necessarily filtered through linguistic categorizations of the phenomena we seek to explain, it becomes clear that there are many alternative descriptions of one and the same reality, and so possibly many alternative explanations of that reality as well. How then can an advocate of the doctrine of explanatory exclusion justify the preference for one level of description as the only one that is legitimate? We herein realize that the principle of explanatory exclusion does not derive directly from supervenience, but instead rests on the acceptance of a prior set of epistemological assumptions that are themselves suspect.

We may conclude that the tension between Kim’s earlier work on supervenience and his later account of explanatory exclusion should create no misgivings over the explanatory agnosticism that is appropriate to supervenience. For the reconciliation between supervenience and explanatory exclusion can only be achieved by adherence to a doctrine that (1) Kim never explicitly states, and (2) even if stated would commit him to a dubious account of the constraints on explanatory adequacy. The philosophical merits of supervenience are many, and its fate deserves to be considered separate from the problematic explanatory account in which it has recently been cast.

REFERENCES


*Received January 26, 2001; revised October 12, 2001; accepted November 7, 2001*