

Mental Causation: The Free Lunch¹

For some time Jaegwon Kim has been telling philosophers of mind that when it comes to mental causation there is “no free lunch.” The “free lunch” he is referring to is served up as accounts of mental causation that allow mental and physical events to causally over determine their effects. The version of the free lunch I like best characterizes mental causation in terms of certain patterns of counterfactual dependencies. Counterfactual accounts are especially attractive to proponents of non-reductive physicalism (NRP) since they permit an effective reply to Kim’s much discussed exclusion argument. The exclusion argument establishes that NRP requires that mental and physical causes, in some sense, “over determine” their effects. Kim thinks such over determination is metaphysically problematic and perhaps even impossible. But there is nothing problematic with an event counterfactually depending both on a mental event and a physical event even when the physical event is a *realization* of the mental event. Kim’s own response to the exclusion argument is a view he characterizes as “physicalism, or something near enough” (Kim 2005). In fact it ends up being part eliminativism and part dualism. It is eliminativism in so far as it denies that there are intentional *properties as such* but only species specific or person specific intentional properties.² It is only “near enough” because, to the surprise of those of his readers who think of Kim as a paradigmatic reductive physicalist, it also ends up being dualism and epiphenomenalism with respect to qualia.³ I think that this is an over reaction to the Exclusion argument. In my contribution I will briefly describe how I understand NRP, why I think it is true, and Kim’s exclusion argument. After these preliminaries I will

¹ I am very grateful to Louise Antony, Karen Bennett, Tim Crane, Carsten Hansen and especially Jaegwon Kim for discussion and comments on earlier versions. Although this paper is critical of his views my debt to him is evident throughout.

² Kim argues that a predicate that is multiply realizable (e.g. a typical functional predicate) does not refer to a genuine property but is reducible instance by instance to whatever physical property realizes it on that particular instance. As we will see this account of reduction is partly motivated by the exclusion argument.

³ The view is proposed in Kim (1998) and is spelled out more extensively in Kim (2005). With regard to phenomenal consciousness Kim says “. . . qualia, are not functionalizable, and hence physically irreducible. . . . There is a possible world just like this world in all respects except for the fact that in that world qualia are distributed differently. (2006 p.170).” I disagree but don’t discuss the claim that the non-functionalizability of qualia predicates/concepts is a good reason for dualism. For discussion see Loar (1997) and Balog (forthcoming).

argue that Kim's exclusion argument- and specifically the claim that over determination is problematic- depends on a particular understanding of causation that physicalists should not find attractive. In contrast counterfactual dependencies between mental and physical events are compatible with physicalism and I will argue are sufficient for mental causation "or something near enough."

I. What is "Non-reductive Physicalism?"

NRP is a metaphysical view of the mind that claims to reconcile physicalism, the irreducibility of mental properties, and mental causation. The way I understand NRP may not be the same as some other writers since I take the "physicalism" part more seriously than some do.⁴ My version of NRP is committed to these claims:

1. Jackson-Lewis Physicalism: Every positive truth and every truth concerning laws and causation is metaphysically necessitated by truths concerning the spatio-temporal distribution of instantiations of fundamental physical properties and relations and the fundamental physical laws.⁵
2. Physical Nomological Closure: For every physical event proposition $E(t')$ (except events at the initial condition) and every prior time t the probability at t of $E(t')$ given the physical state of the universe at t and the fundamental laws is the objective probability at t of $E(t')$.
3. Irreducibility: Some mental properties (events) are real and are not identical to any real physical properties (events).

⁴ Davidson's "anomalous monism" is usually thought of as a version of non-reductive physicalism. But its commitment to physicalism may be rather weak since though Davidson does endorse the causal closure of physics it is not clear that he agrees with the supervenience claim I call "Jackson-Lewis Physicalism" since he denies that there are "tight connections" between the mental and the physical.

⁵ Frank Jackson (1998) and David Chalmers (1996) characterize physicalism as the claim that the every truth is necessitated by totality of truths in the complete language of ideal fundamental physics and the laws of fundamental physics and a statement to the effect that this is the totality of fundamental truths and laws. (The latter can be avoided by restricting the characterization to positive truths). They hold additionally that the entailments required by Physicalism are *a priori*. I do not assume that here. David Lewis (1983) earlier provided a similar characterization of physicalism. There are issues concerning how to define "fundamental physical property or ideal physics" and whether this account is sufficient for physicalism (It is surely necessary). I discuss these issues in Loewer (2001).

4. Mental Causation: Mental properties (events) are causally related (cause) to physical (and other) properties (events).

The first two claims characterize a physicalist world view or what I will call “Physicalism.” Condition (1) expresses the physicalist idea that “all God needed to do” to make the universe is to distribute the fundamental physical properties in space/time and make the laws of fundamental physics. All facts about macroscopic objects, their colors and behaviors and facts about people, their thoughts and experiences, and truths about causation and the special sciences and so on are metaphysically entailed by the fundamental physical facts and laws. Condition (2) says that the physical laws are closed and complete in the sense that given the complete fundamental physical state at t and the laws whether or not E occurs at t' or its chance of occurring is completely determined. I assume that whatever causation is condition (2) implies the *casual* completeness of physics in that $E(t')$'s physical causes at t are sufficient to determine its occurrence (or the chances of its occurrence). Condition (2) is a consequence of (1) and it is possible to derive (1) from (2) and some other plausible premises but I separate them since nomological and causal closure will figure importantly in our discussion.⁶

Physicalism comes in two varieties; reductive physicalism (RP) and non-reductive physicalism (NRP).⁷ RP claims contrary to 3 that every *real* or as I will say “*genuine*” property (G-property) that has instances in our world (or any physically possible world) is identical to a *physical* G- property. NRP claims that some mental properties are G-properties that are not identical to any *physical* G-properties. If events are, as Kim and I think, instantiations of G-properties then there are mental events that are not identical to physical events but are nonetheless *real*. Mental Causation says that some of these mental events cause physical events. It is pretty obvious that NRP is committed to causal overdetermination. Later we will look at how Kim formulates an argument that makes this commitment explicit and attempts to refute it.

⁶ See Papineau (2001), Loewer (1995) and (2001) for arguments from (2) to (1).

⁷ Some philosophers call themselves (or are called by others) “physicalists because they hold that all things are materially constituted even though they reject Physicalism. Perhaps Davidson (19xx) and Searle (19xx) are examples.

The disagreement between advocates of RP and advocates of NRP is an “in house” argument among physicalists. Obviously the issue depends on what counts as a G-property and what counts as a physical property. To explain how I will employ these notions I will adopt a framework devised by David Lewis although without all of his metaphysical commitments.⁸ Lewis calls the conception on which every predicate corresponds to a property the “abundant conception” of properties. The abundant properties are, or correspond to, sets of possible individuals. This is a “thick” notion of property since predicates that differ in meaning may correspond to the same property. For example, “is a puddle of H₂O” and “is a puddle of pure water” differ in meaning but correspond to the same property.⁹ The G-properties are a subset of the abundant properties. Roughly, the idea is that G-properties are those “that cut nature at its joints.” First there are the most fundamental joints. Lewis calls the most fundamental properties “the perfectly natural properties.” Lewis assumes that the space-time distribution all else supervenes *a la* Jackson-Lewis on the distribution of actual fundamental natural properties. Since Lewis thinks Physicalism is true he thinks that all the perfectly natural properties exemplified in our world properties are physical (- *mass, charge, spin*, - these are properties that occur in proposals for the most fundamental laws of physics.¹⁰) Non-physicalist philosophers might want to include properties involving phenomenal consciousness and intentionality; (pain, acquaintance, reference and so on among the fundamental properties.¹¹

⁸ Lewis thinks that any class of possible individuals is a property (the “abundant” conception) and that certain of these classes are, or correspond to, perfectly natural properties (the “sparse” conception) and that naturalness comes in degrees. He also holds that the degree of naturalness of a property is a matter of metaphysical necessity, that the perfectly natural properties instantiated at our world are all intrinsic to space-time points (or small regions) except for space-time relations, and that all truths- including the laws- supervene on the distribution of perfectly natural properties. The latter two comprise his doctrine of Humean Supervenience. I make none of these assumptions in this paper.

⁹ Davidson, whose view Anomalous Monism, is often thought of as a version of NRP was very skeptical about properties but seem to have given in to talking about properties (Davidson 19xx). Needless to say, except in a footnote, the view that predicates with non-analytically connected meanings may correspond to the same property depends on a Fregean-like notion of meaning.

¹⁰ Lewis holds both that physics makes the best estimates of the natural properties and that what properties are natural is a matter of necessity. There is a tension between these commitments.

¹¹ For example, Chalmers (1996) argues for the view that there are fundamental mental or proto-mental properties linked by laws to physical properties. More recently Chalmers (200x) has suggested that mental features might be the categorical basis of fundamental physical properties.

Every perfectly natural property is a G-property but nature also has higher level joints. These correspond to the properties or kinds that occur in laws of special sciences and these too are G-properties¹² By “law” I mean a simple true generalization or equation that is counterfactual supporting, projectible, sufficiently simple and so on.¹³ So the G-properties include the perfectly natural properties and any other properties that are involved in laws.

In order for there to be an interesting difference between RP and NRP there needs to be a restriction on G properties so that not every construction out of physical G-properties is itself a G-property and this restriction needs to be of metaphysical and scientific significance. The distinction between properties that are either fundamental or law involving and those that are not is the distinction that most philosophers have in mind when they speak of some properties as “real” and others as not. Some plausible candidates for G-properties are: *positive charge, being a gas, mutation rate, episodic memory, and being a monetary exchange*. Plausible candidates for not being G-properties are: *being grue, being post-modern, and being a gas or a mutation*, are not G-properties.¹⁴

I will make an assumption about how G-properties figure in laws and causation with which, I think, Kim would agree. It is that if F is a G-property that figures in the antecedent of a dynamical law then an instantiation of F is (or corresponds to) an event and this event can be a cause of other events. So if there is some reason to think that a certain property instantiation cannot be a cause then that is reason to think that it cannot figure in a dynamical law. Perhaps the converse that every causal property occurs in a law (or corresponds to a predicate that occurs in a law) is also true. But this wont figure in my discussion.

¹² Armstrong (19xx) holds this view concerning universals. Of course exactly what this view comes to depends on what laws there are and what it is to enter into a law or causal relation in “an appropriate way.”. Fodor says that *natural kinds* (i.e. genuine properties) are properties that appear in laws and then explains laws by saying that laws are generalizations connecting *natural kinds*. Well, explanation has to end somewhere.

¹³ These are the usual criteria for lawhood. Something along the following lines is what I have in mind. If $F \rightarrow G$ is a law and Fa is logically compatible with its being a law then $Fa \rightarrow Gi$ (G is an appropriate instance of the law) is true and positive instance of $F \rightarrow G$ provide confirmation for further positive instances. Many laws of the special sciences hold *ceteris paribus*.

¹⁴ Of course the instances of any “gruesome” property fall under laws and can be causes. The claim is that the gruesome property does not itself occur in a law or ground a causal relation.

Here is a bit of terminology and some abbreviations that will be useful. A *mental* property, M-property, is any property that corresponds to a mental, that is intentional or a qualia predicate; e.g. “is thinking about soup”, “feels dizzy.” An MG-property is a mental property that is a G-property. A P-property is any property that is picked out by a kind predicate of a natural science. The natural sciences include physics, chemistry, biology, and so on but not intentional/consciousness psychology. Since a disjunction of kind predicates is not necessarily a kind predicate not every broadly physical property is necessarily a P-property. So every P-property is a G-property but it is left open whether M-properties are G-properties and if so are P-properties. There are two important questions for physicalists about M-properties.

1. Are some M-properties MG-properties?
2. Are all MG-properties P-properties?

RP says that if any M-property is a G-property then it is a P-property. NRP says there are M properties that are G properties and that are not identical to any P-property. In other words, RP says that all of nature’s joints are physical while NRP says that some are mental.

There are a number of reasons for thinking that MG-properties are not P-properties. Fodor (1974) observed 30 years ago both that there are apparently laws involving mental predicates and that the property identities that RP claims to exist fail to be found. Aside from this many philosophers are persuaded that certain features of mental properties establish that they are not identical to neurophysiological or any other P-properties. These features are quite familiar to philosophers of mind but a brief discussion of them will be useful as a reminder why they motivate NRP.

The features are a) multiple realizability, b) externalism and c) the existence of an “explanatory gap” between mental and physical facts.

a) Multiple realization: Certain predicates seem to apply not because their instances possess a particular physical constitution but rather in virtue of their instances satisfying a particular causal or functional specification. For example, “is a computer of simple arithmetical functions” applies to a system when it is so configured that when it is in its ready to compute state and is given appropriate representation of numbers and a simple

arithmetical function as input yields a representation of the function's value. It is striking that there are mechanisms that satisfy this general specification that are physically *heterogeneous*. It even seems possible for there to be computers in possible worlds whose physics is very different from the physics of our world. That is, it is possible for there to be computers even though they are made out of *alien* substances following *alien* laws as long as this alien physics implements the causal/nomological profile of a computer. Functionalism about the mind is the view that many psychological predicates are similarly associated with functional specifications.

Despite there being no persuasive *complete* functional analyses of any psychological predicate the view that many psychological predicates denote functional properties became and still is the mainstream view.¹⁵ A functional property is a property that something possesses in virtue of its satisfying a certain causal/nomological/structural specification.¹⁶ If a functional property F is instantiated by X and F is itself not a fundamental property¹⁷ then X satisfies some more fundamental property P (or X is composed of parts that that satisfy more fundamental properties and are configured in a particular way) that satisfies F's causal/nomological profile. P is said to *realize* F.

If a functional property is involved in a law it is a G-property and if it is involved in a natural science law it counts as a P- property. The question is whether mental functional properties are also P-properties (i.e. appear in natural science laws) and it seems that they are not. The reason is that the various possible physical realizers of a mental property are *heterogeneous*. So if there are mental properties that are functional G-properties then RP is false. One response to this would be to broaden the conception of Ps to include configurations of G-property instantiations and arbitrary conjunctions and disjunctions of

¹⁵ There are a number of versions of functionalism. According to analytic functionalism there is an analytic connection between a psychological predicate and a functional specification. Some analytic functionalists (e.g. Jackson XXXX) think that the reference of the predicate is the functional property associated with the specification but others (e.g. Lewis XXXX) think that there are no functional properties as such but the predicate applies if there is a genuine property that satisfies the specification. According to psych-functionalism (e.g. Bloch XXX) there need be no analytic connection between psychological predicates and functional properties but psychological predicates (or certain of them) refer to functional properties.

¹⁶ A different view is that a functional property is a "second order property" - the property of having a first order property that satisfies a certain functional profile.

¹⁷ Some philosophers e.g. Shoemaker (19xx) think that fundamental properties are themselves individuated in terms of their nomological/causal relations. Others e.g. Lewis (1983) think that fundamental properties are *categorical* and it is an entirely contingent matter what laws/causal relations they are involved in. I don't take a stand on this very interesting issue in this paper.

such configurations. But even then functional properties of psychology may not be identical to any physical properties. The reason is that psychological functional properties may be realized by *alien* fundamental properties that conform to alien laws.¹⁸ Suppose the fundamental individuals in our world are atoms or strings. It is at least *prima facie* plausible that there could be a world whose fundamental entities are fields that at the macroscopic level is pretty much indiscernible from the actual world. In this world there are creatures whose behaviors and dispositions to behave are caused in ways that qualify them as having minds even though their physical constitution is very different from ours.

b) Externalism: Putnam's twin earth thought experiments, Burge's arthritis/tharthritis example, Davidson's swamp man and related examples persuaded a substantial part of the philosophical community that intentional states fail to supervene on subject's intrinsic physical condition. The intentional contents of one's beliefs and other attitudes are constituted also by environmental, social, and historical factors. Externalist content constituting features seem to be very complicated and gerrymandered from the point of view of any natural science and so it is quite implausible that any such property is a property that occurs in any natural science law.

c) Explanatory Gap. Joe Levine and others have persuasively argued that there is an *explanatory gap* between the physical description of a person no matter how complete and descriptions of that person's conscious experience employing phenomenal concepts.¹⁹ Even if we were to be sure that pain is perfectly correlated with C-fiber firing it seems that we would have no account of *why* C-fiber firing feels *painful* instead of being associated with different qualia or with none. The situation seems quite different with respect to other macroscopic properties. For example, there is an explanation or explanation sketch, in terms of quantum mechanics of why certain configurations of micro-particles constitute a pane of transparent glass. The difference is that in the latter case it is plausible that there is a functional characterization of what it is to be a pane of

¹⁸ It is not implausible that there to be a worlds whose ultimate constituents are Newtonian particles conforming to Newtonian like laws, worlds whose ultimate constituents are fluids obeying classical fluid mechanics, worlds whose ontology and laws are those of Bohmian quantum mechanics all of which contain configurations that realize the nomological/causal specifications associated with at least some mental properties.

¹⁹ The phrase "explanatory gap" is due to Joe Levine (19xx) but the point that there is a problem with understanding how physical phenomena can constitute consciousness is as old as the philosophy of mind.

transparent glass; i.e. a characterization in terms of allowing the passage of light and so on. If it can be shown that a certain micro-configuration of molecules satisfies the functional specification then we have a constitutive explanation of why that micro-configuration is a transparent pane of glass. But it seems that there is no complete functional specification of phenomenal experience which can play a similar role in accounting for why a certain physical configuration is painful. The point is made vivid by the fact that we can conceive of beings that are functionally and physically identical to us with inverted qualia or with none at all. It is tempting to conclude from this that no physical property or configurations of physical properties instances is identical to or even is metaphysically sufficient for the instantiation of any qualia property.

Before discussing Kim's argument that NRP cannot accommodate mental causation I want to look briefly at two other worries that Kim may have about NRP.

Fodor says that psychological laws are not *reducible* to physical laws. I think this is correct given a particular construal of "reducible." If $L(F,G)$ is a law of psychology then it is not reducible a natural science law iff there is no law of any natural science $L(P,Q)$ that together with true property identities implies $L(F,G)$. But it may seem puzzling how a physicalist can allow that there are irreducible psychological (or anything other fundamental physics) laws since physicalists think that fundamentally everything is the way it is because of physical facts and laws. In the course of defending his view of the relation between the special sciences and physics from Jaegwon Kim's objections Fodor expresses his puzzlement this way:

"So then, *why is there anything except physics?* That, I think, is what is *really* bugging Kim. Well, I admit that I don't know why. I don't even know how to *think about* why. I expect to figure out why there is anything except physics the day before I figure out why there is anything at all, another (and presumably related) metaphysical conundrum that I find perplexing" (Fodor 1998 p.161)

I am not sure that this "is what is *really* bugging Kim" but I do think that Fodor is asking an excellent question. If the laws of physics are complete and closed and everything supervenes on the physical how is there room for additional laws? How can there be laws other than the laws of physics? At places Fodor seems to suggest that for there to be

special science laws there must be more in the world than can in principle be accounted for (in principle) by physics; that the fact that certain special science generalizations *are laws* is a fact that does fail to supervene on the physical laws and facts.²⁰ But if Physicalism is true then if it is a law that Fs are followed by Gs cp the laws of physics *together* with additional *solely* physical facts entails that it is a law that Fs are followed by Gs cp. The story of what additional facts are needed and why it is plausible that they do entail special science laws that are not reducible to laws of physics (in the sense of reduction at issue in the dispute between RP and NRP) is a complex issue that I address elsewhere.²¹ But even without an answer to Fodor's question one can see that there is no contradiction between Physicalism and the existence of laws involving non-physical mental properties. We can see this because we know that there are laws of e.g. meteorology that are *ceteris paribus laws* and that the dynamical laws of physics are exception less and so cannot together with property identities entail meteorological generalizations that hold. But it is absurd to think that the law hood of meteorological laws doesn't supervene on the fundamental physical laws and facts.

There is another line of thought (suggested by some of Kim's discussion and briefly adverted to above) that threatens to show that NRP is unstable and make my discussion of his exclusion argument (and the argument itself as an argument against NRP) moot. Suppose that M is a mental property and occurs in some law, say L(M,R) (the law may be a *ceteris paribus law*) and so is a G-property. Suppose also that Physicalism is true. Won't there be some property Q constructed out of P-properties that is co-extensive with M in all physically possible worlds? But then won't it be the case that L(Q,R) is also a law? At least this follows if we can substitute terms that refer to physically co-extensive properties in laws. If so it follows that Q is a G-property since it figures in a law. And since Q is constructed out of physical properties it too is physical and so is a P-property. If this is correct then NRP comes very close to collapsing into RP since either

²⁰ Fodor seems to me to be pretty cagey with respect to this issue. In "The Disunity of the Sciences..." he says that while lower level physical laws and realization statements entail the higher level generalization they don't entail that it is a law. This sounds like the view I am attributing to him. On the other hand, in other places he seems to espouse a thoroughgoing physicalism e.g. "If intentionality is real it must really be something else" (Fodor 199x).

²¹ See Loewer (2006) "Why There is Anything Except Physics"

$M=Q$ or $M^*=Q$ where M^* is the property M restricted to the class of physically possible worlds. In the first case RP holds in the second case it is close enough to make the difference between RP and NRP look awfully trivial.²²

There are two responses to this argument. One is to say that that if any construct out of physical properties that is co-extensive (or co-extensive in every physically possible world) with a G property counts as a P -property then indeed NRP and RP come to much the same. But as a defense of RP this maneuver looks a lot like “declaring victory and withdrawing” since it amounts to claiming that R is a G -property in virtue of satisfying a *psychological* law. The second response is to deny that $L(Q,R)$ is a law. If laws are thought of as generalizations or equations (i.e. as *proposition* like) then this is a natural response. Fodor points out (Fodor 1999) discussing a point similar to this one that “__ is a law” is an intensional context. This is so on accounts of laws on which a condition on law hood is that laws are instance confirmable since confirmation is an epistemic notion.²³ If confirmation is understood in subjective Bayesian terms then this point is obvious. A subjective probability distribution on which $M \rightarrow R$ is confirmable but $Q \rightarrow R$ is not is perfectly coherent and given the complexity of R is quite reasonable. I conclude that this line of thought doesn’t work to undermine NRP or render the difference between NRP and RP uninteresting. Let’s see if the exclusion argument is more successful.

II. The Supervenience/Exclusion Argument:

²² Karen Bennett (200x) also makes this point.

²³ If confirmation is understood in Bayesian terms then this point is obvious. A subjective probability distribution on which $M \rightarrow R$ is confirmable but $Q \rightarrow R$ is not is perfectly coherent and given the complexity of R is quite reasonable.

Kim's supervenience/exclusion argument has received a lot of discussion so I will be going over some well (in some cases very *well*) tread ground. Here is how Kim recently formulates the exclusion argument (2004 p.39)

Let M and M^* be mental properties and m and m be the events of M 's instantiation at some location and time t and m^* be the event of M^* 's instantiation at some place and time t^* and suppose that

(1) m causes m^* ²⁴

In saying that m causes m^* Kim is supposing that m is causally sufficient in the circumstances for m^* . Since physicalism holds there will be some physical property P^* whose instantiation p^* at time t^* is such that

(2) m^* has p^* as its supervenience base²⁵

Kim argues that (1) and (2) support

(3) m caused m^* by causing p^*

Since physicalism holds

(4) m also has a physical supervenience base p .

Kim then appeals to a principle he calls "Closure" (what we earlier called "the causal completeness of physics")

Closure: If a physical event has a cause that occurs at t , it has a physical cause that occurs at t .

It follows that

(5) m causes p^* and p causes p^*

Since we are assuming NRP i.e. non identity

(6) $M \neq P$ and so $m \neq p$

At this point in the argument Kim appeals to a principle that he calls "Exclusion."

Exclusion: No single event can have more than *one sufficient* (my ital) cause occurring at any given time—unless it is a case of causal over determination.

²⁴ Kim talks of properties being in causal relations and also property instances being in causal relations. He identifies events with property instances so the latter involves event causation.

²⁵ That is there is some physical fact that is metaphysically sufficient for M^* . We can think of P^* as the property of this fact obtaining at some region of space-time. Although Kim doesn't emphasize the point P^* may be enormously complicated and may involve events in a temporal region. It may not be a *genuine* physical property.

But according to Kim this isn't a case of causal over determination. By causal over determination Kim means the kind of case in which there are two shooters each of which kills the victim. That seems right.

(7) p^* is not causally over determined by m and p

It follows that either p or m does not cause p^* . By closure it must be that

(8) The putative mental cause m is excluded by the physical cause p . That is, p not m , is a cause of p .

As Kim observes Supervenience isn't needed for the argument. The conflict is among $M \neq P$, Closure, and Exclusion. It appears then that the argument works equally well against NRP and against non-physicalist views.²⁶

III

The Exclusion Argument Defanged

At this point it may be useful to remind ourselves what is at stake in the exclusion argument. Various considerations (functionalism, multiple realization, externalism, explanatory gap) make $M \neq P$ persuasive. Also there are scientifically compelling reasons to accept Physicalism and Closure. Giving up mental causation is a last resort. So Kim's argument is a paradox. Each of $M \neq P$, Closure, Mental Causation, and Exclusion is plausible but together they are inconsistent. We proponents of NRP accept $M \neq P$, Closure, and Mental Causation so we have to reject Exclusion.

One response to the Exclusion argument is that it must be unsound since the parallel argument in which P is restricted to fundamental physical properties and M is any multiply realized special science property would show either that M is reducible to fundamental physical properties or that M isn't causal.²⁷ Kim calls this "the generalization argument" and attempts to rebut it. I don't intend to get into the details of his reply since my primary response to the argument will be to attack Exclusion directly. However, I do want to discuss one way in which Kim responds to the generalization

²⁶ Actually, as I will argue in the next section the argument doesn't work against NRP but has some bite against non-physicalist emergentism.

²⁷ The Generalization Argument has been made by Gillett (2001), Hansen (200x), Block (2001) among others.

argument since it provides insight into how he is thinking about causation that will be relevant to my criticisms of Exclusion. The heart of his response involves a distinction he makes between levels and orders and his view that causation is grounded in *causal powers*. At the most fundamental level there are elementary particles and at the next level are certain configurations of elementary particles and so one up a ladder of levels. On Kim's account a property P may bestow certain causal powers on an individual X that are not had by any of X's lower level components. For example, a brain possesses causal powers in virtue of being a brain that are not possessed by its component neurons and neurons have causal powers that are not possessed by their molecular components and so on down to the most elementary particles. Kim's suggests that some special science properties apply to higher level individuals and bestow causal powers not bestowed by lower level properties. A functional property and its realizer properties apply to individuals at the same level but of different orders. The functional property is second order since it applies in virtue of an individual possessing some first order property that realizes it. On this picture the functional property doesn't contribute any new causal powers over and above its realizer to the individuals that instantiate them. Since Kim supposes that certain psychological properties (although not consciousness properties) are second order and apply at the same level as brain properties they are subject to the exclusion argument. But those special science properties that apply at higher levels and are first order may contribute novel causal powers and so are not subject to the exclusion argument. I think that is basically the idea.

Kim's attempt to save some macro properties from the Exclusion argument depends on this metaphysics of level and orders and on his thinking of causation in terms of the causal powers properties bestow on individuals. I have a lot of doubts about this metaphysical framework.²⁸ As we will see it doesn't fit well with the way causation is conceived in fundamental physics. In any case it doesn't really provide an adequate defense against the generalization argument. Many special science properties other than psychological properties are functional properties that apply at the same level as their

²⁸ I am not at all sure how to assign levels to configurations of particles. Are water molecules and sugar molecules at the same or different or incomparable levels? What if fundamental physics says (as it likely does) that what basically exists are fields of various kinds? Kim discusses some of the problems with "the layered" account in Kim (2002).

realizers and so are just as subject to the exclusion argument as psychological properties. Further, it seems that the only properties that escape the Exclusion argument, if it is sound, are the microphysical states of isolated systems. For example, being a low pressure system is not identical to any particular arrangement of particles and fields since there are infinitely many distinct micro states that make for a low pressure system but at best is necessarily a disjunction of these states. Nomological closure/completeness implies that whatever causal powers a specific low pressure system has it has in virtue of the fundamental physical laws applied to its fundamental physical state. Speaking the language of “causal powers” it appears that being a low pressure system doesn’t contribute any *new* causal powers over and above the causal powers of the microphysical state. Anyway, that’s why I don’t find Kim’s reply to the generalization argument effective. But, as Kim points out, the generalization argument at best shows that *something* is wrong with the exclusion argument (or that it threatens a lot more than we might have thought) it doesn’t tell us *what* is wrong explain *why* it is wrong. A better response to the exclusion argument is to show why Exclusion is wrong. So let’s examine this premise more carefully.

Exclusion says that “no single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of causal over determination.” By “a genuine case of causal over determination” Kim has in mind the type of situation in which two assassins fire simultaneously at the victim causing her death. In this kind of situation the two events (the two firings) are *metaphysically* independent and each involves its own causal process that culminates in the death of the victim. Causal over determination like this may be rare but it is not metaphysically problematic. Kim observes that the putative situation in which a non-physical genuine property instantiation $M(y,t)$ and its physical realizer $P(y,t)$ are said to both cause $Q(z,t')$ is not a case of genuine causal over determination like this. He is certainly correct about this since $P(y,t)$ and $M(y,t)$ are not metaphysically independent. Let’s call the kind of over determination involved in mental/physical causation “M-over determination” and understand Kim as ruling it out by the Exclusion principle. According to Exclusion the putative mental and physical causes of $Q(z,t')$ compete and so one is not really a cause of $Q(z,t')$. Since NRP

assumes that the physical realm is causally closed $P(y,t)$ wins the competition and $M(y,t)$ is not a cause of $Q(z,t')$.

The exclusion argument assumes that the physical cause and the putative mental cause are *sufficient* causes of $Q(z,t')$. However, If $M(y,t)$ and $P(y,t)$ are ordinary macro events then, contrary to Kim's supposition, they are certainly not by themselves sufficient for $Q(z,t')$. It is a common place among philosophers of science, but perhaps not as recognized as it should be outside of philosophy of science, that for any small region R of space at time t nothing much short of the state of the universe in a sphere with center R and whose radius is one light second (i.e. 186,000 miles) at $t-1$ second is causally sufficient for determining what will occur (or the chances at $t-1$ of what will occur) in R .²⁹ Because of this I suggest that we interpret Kim's exclusion principle not as involving causes that are literally sufficient for their effects but as nomologically sufficient in the circumstances C where C is a partial description of the state at t .

Exclusion*: there can't be two distinct events $P(x,t)$ and $F(x,t)$ such that both are causes of $Q(z,t')$ and there are circumstances C such that both $P(x,t)$ and $F(x,t)$ are nomologically sufficient for $Q(z,t')$ in C . for $Q(z,t')$ (or are each nomologically sufficient in C for the chance at t of $Q(z,t')$).

Kim doesn't argue for Exclusion or Exclusion* since he thinks that it is "is virtually an analytic truth without much content" (p.51). This is puzzling since it is sufficiently contentful to play an essential role in the argument that NRP is incoherent and philosophers who deny it, myself among them, don't think that we are denying an analytic truth. So there must be some assumptions about the nature of causation that Kim accepts and that lead him to think that M-over determination is "virtually analytic". I don't know exactly what he is assuming but I want to offer a speculative hypothesis.

²⁹ Or if the fundamental laws are deterministic as determining the chance at t of $Q(z,t')$. More exactly, for any event E at t' there will be a physical proposition K that holds at time t that is a minimal sufficient condition for the occurrence of E (given the physical laws) which is typically a partial description of the complete state at t (or state on a hyper surface intersecting t) but this proposition will involve values of physical parameters throughout the hyper surface. This point is made by Latham, Field, Loewer, and Elga among others.

Ned Hall (2004) has recently argued that there are two concepts of causation which he calls “production” and “dependence.” Production is the relation that supposedly obtains when one billiard ball hits another and thus *produces* motion in the second. Dependence is the relation that holds between two events when features of the second (including whether or not it occurs) counterfactually depends on features of the first (including whether or not it occurs). Hall thinks that the two kinds often go together but are fundamentally different and that it is possible to have one without the other. For example, the kitchen fire may depend on my forgetting to turn off the heat under a pot but my forgetting does not *produce* the fire. Billy’s throw may produce the breaking window even though, in this instance, *dependence* is absent since Billy’s rock arrived a moment before Sally’s which would have broken the window anyway.

My diagnosis of why Kim thinks that Exclusion is virtually analytic is that he is thinking of causation as *production*. If one thinks about causation in this way then it is quite natural to see Exclusion as virtually analytic. If $P(x,t)$ literally produces $Q(z,t')$ then it does appear that “there is no work” left for any other event $F(x,t)$ to do as far as producing $Q(z,t')$. Kim also seems to think of the causal relation as involving the transfer of some quantity, *causal oomph*, from the cause that brings the effect into existence. It is not surprising then that he would think that a second dose of *oomph* from $F(x,t)$ is not only not needed to produce $Q(z,t')$ there isn’t even any place for it.

I am not confident that Kim is thinking of causation in this way but I am sure that no one who accepts Physicalism should endorse it. In a famous essay Bertrand Russell noted that causation as a relation of production connecting local events makes no appearance in the ontology of the fundamental laws of physics and suggested that the idea of causation “...is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm.”³⁰ The problem he finds with it is that while productive causation connects localized events at different times the candidates for

³⁰ “All philosophers, of every school, imagine that **causation** is one of the fundamental axioms or postulates of science, yet, oddly enough, in advanced sciences such as gravitational astronomy, the word “cause” never appears. Dr. James Ward... makes this a ground of complaint against physics... To me, it seems that... the reason why physics has ceased to look for causes is that, in fact, there are no such things. ... The law of causality, I believe, like much that passes muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm. “

fundamental laws are differential equations that connect global states with one another. If one wants to speak of “production” it seems that it is the whole state and the laws that produce subsequent states. Russell’s suggestion that we would be better off without the concept of causation, as he also came to believe, is an over reaction. The appropriate response for a Physicalist is to characterize causal relations that supervene on the fundamental laws and facts. I don’t know whether there is account of causation as production that can support Exclusion and is also compatible with Physicalism. But even if there is all it would establish is that NRP is not compatible with mental events being productive causes so understood. This conclusion would have the import that Kim thinks his Exclusion argument has only if it is causation as production (construed in this way) that we think is really required to vindicate mental causation. But I think a good case can be made that causation as dependence will do perfectly well.

Causation as Dependence

Counterfactuals are notoriously vague and context dependent. The way they should be understood for the purposes of characterizing causation as dependence is along the lines (but not quite; see fn. 32) of David Lewis’ famous account. On that account $A > B$ is true if either there are worlds at which $A \& B$ are true that are more similar to the actual world than any world at which $A \& \neg B$ is true. Lewis specifies a particular account of world similarity that he thinks has the consequence that in evaluating $A > B$ one looks at worlds that are identical to the actual world from the worlds initial condition and then diverges from the actual world (perhaps this requires a violation of actual laws in small region for a short time) and then evolves in accord with the laws of the actual world so that A is true. If all these worlds are all worlds at which B is true then $A > B$ is true. For example “if at noon Terry had wanted a beer he would have opened the refrigerator” is

true if the worlds that are identical to the actual world up until a moment prior to noon when a small miracle occurs so that Terry is in a brain state that realizes wanting a beer are also worlds where he opens the refrigerator.³¹

Lewis says that E depends on C iff C and E are non-overlapping events and if C had not occurred the E would not have occurred. His original account of causation was that C causes E iff C and E occur and E depends on C or there is a chain of events connected by dependence from C to E. This account is vulnerable to cases of preemption in which C causes E but E doesn't depend on C because there is another event C* waiting in the wings to cause E if C didn't occur. Lewis modified his account in a way that handles many preemption counterexamples. On the most recent account an event E *influences* an event C if E and C don't overlap and if there are suitable variations in C that are counterfactually correlated with variations in E. C causes (in the dependence sense) E iff C and E occur and there is a chain of events connected by influence from C to E. For example, the height of mercury in a thermometer depends on the ambient temperature since the counterfactuals "if the temperature had had been (or were to be) x the height of the thermometer would have been (would be) y" are true for a range of x and y.

It is plausible that under normal conditions small differences in a person's brain corresponding to different mental states (e.g. different intentions) lead by law to correspondingly different bodily movements. That is, that counterfactual dependencies *on Lewis' construal of counterfactuals* between mental events and bodily events obtains. If so then mental events cause *in the dependence sense* bodily events. My proposal is not that Lewis' influence account perfectly captures our intuitive concept of causation. But I do claim that causation as influence is near enough to our folk conception of mental causation to underwrite the role of causation in folk psychology, rational deliberation, action theory and so on. In the remainder of this paper I will lay out a case for this claim.

³¹ Unfortunately, Lewis' account of world similarity doesn't have the consequence he thinks it has. The heart of the problem is that his account of similarity involves laws and other considerations that are temporally symmetrical while the similarity he thinks the gets out of these considerations is temporally asymmetric, as it must be if it is going to get the truth values connected with causation as dependence correct since these counterfactuals are temporally asymmetric. See Elga (20xx) Jonathan Bennett (200x) characterizes truth conditions of counterfactuals by simply counting past perfect match and not future march as making for similarity. It is possible to fix this all up by adding a bit to Lewis account so that one gets more or less the similarity relation Lewis was aiming at but it would take us too far a field to do it here. See Loewer (2006a) for the fix up.

The first thing to note is that there is no problem of over determination if causation is understood as dependence. On Lewis' account of counterfactuals a particular event (or the value of a range of possible events) can depend on many co-occurring events. The motions of one's body, for example, the motions of a person's arms and hands when reaching into the refrigerator depends counterfactually both on her mental states (which snacks she wants) and on her brain (and other bodily) states and on a myriad of other states and events. Also, the kind of "M-over determination" involved in B depending on both M and P is neither like the two assassins kind nor the production kind. In particular there is no temptation to say that if B depends on P it can't also depend on M since "there is no work for M to do." If there is "work being done" it is being done by the *fundamental* dynamical laws that evolve the entire state. The influence counterfactuals connect aspects of the state at one time to aspects at other times so that alterations of one correlate with alteration of the other.

Kim expresses his worries about counterfactual accounts in this passage:

"To summarize our discussion of the counterfactual approach then, what the counterfactual theorists need to do is to give an *account* of just what makes those mind-body counterfactuals we want for mental causation true and show that on that account those counterfactuals we don't want, for example epiphenomenalist counterfactuals, turn out to be false. Merely to point to the apparent truth, and acceptability of certain mind-body counterfactuals as a vindication of mind body causation is to misconstrue the philosophical task at hand.what we want- at least what some of us are looking for- is a philosophical account of *how* it (mental causation and the corresponding counterfactuals) can be real in light of other principles and truths that seem to be forced upon us."

I agree with Kim that merely pointing out that certain counterfactuals are true or appear to be true is not sufficient to ground mental causation. As he says it is also required for there to be an account of how those counterfactuals come to be true and further that account cannot presuppose mental causation and must be compatible with NRP. Lewis' account does this. On Lewis account the fundamental laws and facts of physics is "what makes those mind-body counterfactuals we want for mental causation" true.³² "If I were to decide to get a beer I would walk over to the refrigerator" (and

³² Kim suggests ("Causation and Mental Causation) that the truth makers of counterfactuals or the counterfactuals that go along with mental causation involve causation as production. This is correct if one has in mind the fundamental physical laws evolving fundamental physical states. But Kim is more likely thinking of what I called "local production." Relations of local production are not the truth makers of

similarly for the battery of other counterfactuals that ground causation as dependence) is true when the worlds most similar to the actual world in which I decide to get a beer are worlds in which I walk over to the refrigerator. Whether that is so depends on the actual laws of physics (since what they are determines what counts as a “small violation”) and on the actual physical facts. It is clear that this account of counterfactuals does not presuppose causation or mental causation and is compatible with physicalism. Also, we have seen that there is no problem about over determination so the account is compatible with $M \neq P$ i.e. with NRP. So the issue that remains whether the account of counterfactuals really underwrites “those counterfactuals we want” and not “those we don’t want.” Fully establishing these claims is not something that I can do since it would involve establishing the truth values of many counterfactuals (on Lewis’ construal) and that can literally be done only by knowing the physical realizers of mental states and the fundamental laws. But I think I can go some distance towards making the claim plausible and replying to Kim’s arguments that causation as dependence cannot do the work we want mental causation to do.

Kim suggests that there may be dependence/influence where there is no mind-body causation. If so then dependence is too weak to ground genuine mind body causation. He mentions four kinds of situations; backtrackers, common causes, epiphenomenalism and omissions, where dependence holds but there is no causation.

Although Kim doesn’t discuss backtracking in detail I think he may have something like the following worry in mind. On Lewis’ account the worlds relevant to evaluating $A(t) > B$ are ones in which at the nearest time prior to t there is a small violation of law that leads to $A(t)$. But for some $A(t)$ that time may be much prior to t . For example, had Haley’s comet intersected the orbit of Jupiter at t it would have had to intersect the orbit of Saturn at $t-k$. Getting the comet intersecting Jupiter’s orbit at t when in fact the comet is near Saturn at t would involve too big a miracle at times after $t-k$.³³ For the planetary system it may be that the past counterfactually depends on the present as much as the future does. This is part of the reason that I don’t think Lewis’ account by

counterfactuals on Lewis’s account. The fundamental laws and fundamental physical state are the ultimate truth makes of both kinds of causal relations.

³³ Note that the point isn’t that we are evaluating the counterfactual in what Lewis thinks of as a backtracking way of evaluating similarity but rather that Lewis’ way of evaluating similarity leads to backtracking in certain situations.

itself corresponds exactly to our notion of causation. One would have to add some further conditions (perhaps that under usual conditions causation is counterfactual dependence from past to future). But this is not a problem for mental causation because various decisions, intentions, and so on correspond to very small differences in the brain. Partly for this reason it will always be possible for the violation of law that is required for a counterfactual mental event to occur almost immediately prior to the mental event. I will return to this point later.

C is a common cause of A and B when C causes both A and B but there is no causal relation between A and B. For example a rock thrown into the center of a pool (C) causes a wave to hit at point *a* and at point *b* at time *t*. The worry is that the counterfactual “if A had not occurred then B would not have occurred” may appear to be true. In fact I think that in ordinary language this counterfactual is plausibly true in the situation I described. But recall that the characterization of causation as dependence involves a very particular way of evaluating counterfactuals. On that way this counterfactual is false since the world in which a small violation of law occurs just before *t* that leads to A not occurring but leaves all else the same including B is more similar world to the actual world than the world that also leads to the wave not hitting *b* at *t*. Again as in the backtracking case there may be systems that are so set up so that one does obtain counterfactual dependence between events that are effects of a common cause. But this won't occur with respect to mental events and their putative effects.

An interesting example of a possible common cause situation has come up in the philosophy of mind concerning the relationship between the conscious decision to act in a certain way. There is evidence that at least in some cases the decision and the act are related as the common effects of a brain event that is the common cause of both.³⁴ Whether or not this is so it is clear that causation as dependence/influence has no trouble distinguishing between the decision being the cause of the act or the two being common causes of an unconscious brain event.

Kim raises another worry about dependence that is related to the common cause objection. He argues that causal dependence cannot distinguish the situation in which mental events are genuine causes from the view in which they are mere epiphenomena

³⁴ Daniel R. Wegner, D.M. (2002)

that are nomologically correlated with brain events that are the genuine causes. Kim pictures the situation involving mental causation as follows:

Kim's Favorite Diagram



An epiphenomenalist like T.H. Huxely holds that P1 and P2 are events that are sufficient-in senses to be specified- respectively for the events M1 and M2 and that there is a genuine causal relation between P1 and P2 but not between M1 and P2 (or M2). Kim claims that this is completely compatible with both P1 and M1 causing in the dependence P2. How one should respond to Kim depends on how he is thinking of the strength of the relation between the Ps and the Ms depicted by the *vertical* lines. Epiphenomenalists generally think of this relation as *weaker* than metaphysical necessitation. Perhaps it is nomological.³⁵ Kim likes to illustrate epiphenomenalism with the example of the positions of a shadow cast by a moving ball that *seem* to be causally connected. The positions of the shadow are nomologically connected to the positions of the ball that casts the shadow but are not causally related to each other. Kim seems to think that the counterfactual account fails to count this as epiphenomenalism since $\neg M1 > \neg P2$ will be true. But if counterfactuals are evaluated along Lewisian lines it is not clear that we obtain this result. It is plausible that $\neg M1 > \neg P2$ fails since the most similar world in which $\neg M1$ holds is one in which the vertical law connecting P1 to M1 is broken while the horizontal law connecting P1 to P2 continues to hold.³⁶ On the other hand, $\neg P1 > \neg P2$ may be true. In contrast to this NRP holds that the connection between P1 and M1 is one of

³⁵ Chalmers (1996) suggests that qualia are connected by law to physical systems and are epiphenomenal.

³⁶ Of course there are contexts in which the counterfactual "if the shadow had not been at position x at time t the ball would not have been at position y at time t+" but it is important to keep in mind that the relevant account is Lewis'. On that account the counterfactual comes is false since small violations in law that changes the position of the shadow leave the position of the ball as it was.

metaphysical not merely nomological necessitation. In the most similar worlds at which – M1 it is also –P1 since there is no question of “breaking” the metaphysical connection. So in this situation –M1>-P2 may well be true. But it would be question begging to say that M1 isn’t *really* a cause P2 in this case say because it doesn’t produce or transfer *oomph* to P2.

The last problem is that Kim points out that dependence can connect omissions with events. Kim says:

Friends of the counterfactual approach often tout its ability to handle omissions and absences as causes and the productive/generative approach’s inability to account for them. We are inclined to take the truth of a counterfactual like:

If Mary had watered my plants, the plants would not have died

as showing that Mary’s not watering, an omission, caused the plants’ death and take that as a basis for blaming Mary for killing the plants. But obviously there was no flow of energy from Mary to the plants during my absence (that exactly was the problem!); nor was there any other physical connection, or any spatiotemporally contiguous chain of causally connected events.

Kim’s objection seems to be that since dependence can connect an omission (Mary’s not watering the plants) with an event (the plant’s dying) even though there is no transfer of energy from Mary to the plants dependence cannot really be what we want by mental causation. He says of it “This is not causation worth having.” But, in the first place unless Mary is outside of the back light cone of the plant’s death there will almost certainly be some energy transferred from her to the plants; just not in the right way to save the plants. In any case, omissions are not events it doesn’t follow from the fact that there is dependence on omissions that dependence on commissions, and specifically the counterfactual sensitivity of the positions of one’s body (and, fingers, and so on) to one’s volitions and the counterfactual sensitivity of one’s volitions on one’s intentions, beliefs (and so on) is “not causation worth having.” Indeed, these relations of dependence and influence are absolutely essential to mentality and action. If the transfer of energy is involved in any case of genuine mental causation it is also likely involved in any case of mental causation as influence. But the mere transfer of energy certainly isn’t what we want for mental causation! My conclusion is that Kim has not shown that counterfactual

dependence (underwritten by an account of counterfactuals along Lewisian lines) is not sufficient for genuine mental causation.

want to conclude with a few sketchy and perhaps surprising remarks about the connection between Lewisian accounts of counterfactuals and mental causation. I have appealed to Lewis' account of causation as dependence to ground mental causation. But there is a way in which mental causation, or more precisely, our neural/cognitive structure, also grounds Lewis' account of counterfactuals and thus causation as dependence.

On Lewis' account the candidates for most similar worlds in which the counterfactual antecedent $A(t)$ is true are those whose pasts match the actual past until a short (or as short as can be) time prior to t and then diverges by a small local violation of law and then evolves in accordance with the actual laws. But why, we may ask are we interested in *this* notion of similarity among the infinity of possible similarity relations that satisfy Lewis semantics for conditionals?³⁷ One might think that the answer is that this relation is or at any rate is close to tracking the causal relation and we are interested in that relation because it is a fundamental relation between events. But I think this has things backwards. My view is that we are interested in the causal relation not because it is a fundamental relation – there is no fundamental causal relation to be found in physics that connects local events in the way causation is alleged to - but rather because Lewis' account tracks our ability to influence the likelihoods of events.. Here is what I have in mind. We assume that the alternative decisions that we might make in the next few moments correspond to very small local physical differences in our brains. That is, different decisions that one might make are realized in differential brain phenomena that can result via the laws from tiny microscopic differences immediately prior to the decisions.. If the laws are deterministic then these small differences from actuality involve small localized violations of law. If the laws are indeterministic then the alternative decisions can be reached by chance.

Naturally we are interested in what will happen on the alternative hypotheses of each decision. Of course, that depends not only on the decision but also on many other matters in the environment. For example, suppose Nixon is deciding whether to press the button

³⁷ This question is asked by Horwich (1987) and answered more fully than by suggestion here in Albert and Loewer (2006)

marked marked “Launch.” Assuming that Nixon’s body, hand, fingers and so on are appropriately connected to his brain then what will happen depends on the buttons being hooked up to a various further devices. The interesting point for us is that what will happen, or if we allow probabilities over micro histories, the probabilities of what will happen are given by adding one or the other decision- or rather the physical phenomena that realize them- to the state that is most similar to the actual state that contains the brain state corresponding to the decision. So the reason we are interested in evaluating counterfactuals along Lewisian lines (or rather along the lines that he thought his proposal yields and my amended account does yield as mentioned in footnote 30) is that conditionals so evaluated contain information about the likely results of our decisions and this information is enormously important to our getting what we want. If this is one the right track then causation as dependence has its origin and is most at home in “the actions of our will;” that is in mental causation. Counterfactual dependence evaluated in terms of the Lewisian account of similarity (or rather my amended account) can be seen as a generalization from the decision situation. The worlds that count as most similar are those that match the actual world until a short (as short as possible) time before the antecedent and then lead by a tiny local violation of law (of the sort required for alternative decisions) to the antecedent. Causation as dependence is then characterized in terms of this counterfactual. And causation as dependence *is* causation, or “something near enough” to be genuine mental causation.

Barry Loewer
Philosophy Department
Rutgers University
March 1, 2006

References:

Armstrong, David (1978) *A Theory of Universals vol.2* Cambridge University Press

Balog, Katalin (forthcoming) "Mental Quotation"

Bennett, Jonathan (2001) *Textbook on Conditionals*

Bennett, Karen (2003) "Why the Exclusion Problem Seems Intractable, and How, Just Maybe, To Tract It." *Nous* 37:3, pp. 471-497 (September 2003)

Bennett, Karen (forthcoming) "Exclusion Again"

Block, Ned (1994) "Functionalism" in *Blackwell's Companion to the Philosophy of Mind* ed. S.Guttrnplan Blackwells: Oxford

Block, Ned (2003) "Do Casual Powers Drain Away?" *Philosophy and Phenomenological Research* 67; 2003 pp. 133-150

Chalmers, David (1996) *The Conscious Mind*, Oxford: Oxford University Press

Crane, Tim (1995) '[The mental causation debate](#)' *Proceedings of the Aristotelian Society, Supplementary Volume* 69, 1995, 211-236

Crisp, Thomas and Ted Warfield (2001). "Kim's Master Argument," *Nous* 35 (2001):304-16

Davidson, Donald (1980) "Mental Events" reprinted in his *Essays on Actions and Events* Oxford and New York: Oxford University Press.

Elga, Adam (2001) "Statistical Mechanics and the Asymmetry of Counterfactual Dependence," *Philosophy of Science* 68 Proceedings, 2001.

Elga, Adam (2006) "Isolation and folk causation", forthcoming in *Russell's Republic: The Place of Causation in the Constitution of Reality*. Huw Price and Richard Corry, eds. Oxford University Press.

Field, Hartry. (2003) "Causation in a Physical World", in M. Loux and D. Zimmerman, eds., *Oxford Handbook of Metaphysics*, Oxford University Press 2003.

Fodor, Jerry (1974) *The Language of Thought* M.I.T. Press

Fodor, Jerry (1992) *Psychosemantics* M.I.T. Press

Gillett, Carl (2001) "Does the Argument from Generalization Generalize? Responses to Kim," *Southern Journal of Philosophy* 39; 2001

Gillett, Carl and Loewer, Barry (2001) *Physicalism and its Discontents* Cambridge University Press.

Hall, Ned (2004) "Two Concepts of Causation" in *Causation and Counterfactuals* ed. Collins, J. , Hall, N. , and Paul, L. M.I.T. Press pp. 225-276

Hansen, Carsten (2000) "Between a Rock and a Hard Place" *Inquiry*

Horgan, Terence (1993). "Nonreductive Physicalism and the Explanatory Autonomy of Psychology," in *Naturalism: A Critical Appraisal* ed. Wagner and Warner (Notre Dame, IN: Notre Dame University Press 1993)

Horwich, Paul. (1987) *Asymmetries in Time*. Cambridge, Mass, MIT Press

Jackson, Frank (1998) *From Metaphysics to Ethics: A defense of conceptual analysis*, Oxford: Oxford University Press

Kim, Jaegwon (1998). *Mind in a Physical World* (Cambridge: Cambridge University Press).

Kim, Jaegwon (2002) "The Layered Model: Metaphysical Considerations," *Philosophical Explorations* 5 (2002) 2-20.

Kim, Jaegwon, (2005) *Physicalism: Or Something Near Enough* Princeton University Press.

Kim, Jaegwon, (2006) "Causation and Mental Casuation" This volume

Lewis, David (1986) *Collected Papers* Vol II Oxford University Press

LePore, Ernest and Loewer, Barry (1987) "Making Mind Matter", *The Journal of Philosophy* 93: 63-42.

LePore, Ernest and Loewer, Barry "More on Making Mind Matter More"

Loar, Brian (1997) "Phenomenal States: Second Version", in *Consciousness*, ed. Block, Flanagan, Guzeldier, MIT Press, 1997.

Loewer, Barry (1995) "An Argument for Strong Supervenience" in [Savellos, E./Yalçin, Ü.D.](#), eds. 1995. *Supervenience: New Essays*. Cambridge, MA: Cambridge University Press.

Loewer, Barry "Review of Jaegwon Kim's *Mind in a Physical World* in *The Journal of Philosophy* 98, 315-24.

Loewer, Barry (2002) "Comments on Jaegwon Kim's *Mind in a Physical World*" *Philosophy and Phenomenological Research* 65:655-62

Loewer, Barry (2001) "From Physics to Physicalism" in *Physicalism and its Discontents* ed. Gillett, C., and Loewer, B. Cambridge University Press.

Loewer, Barry (2006a) "Counterfactuals and the Second Law" *Causal Republicanism* ed. Huw Price and Richard Corry. Oxford: Oxford University Press

Loewer, Barry (2006b) "Why There is Anything Except Physics" in volume on the Philosophy of Jerry Fodor ed. Tim Crane. Oxford: Oxford University Press

Mclaughlin Brian (1989) "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical," *Philosophical Perspectives*, 3, (1989):109-34.

Melnyk, Andrew (2003) *A Physicalist Manifesto* Cambridge University Press

Melnyk, Andrew Review of Jaegwon Kim's *Physicalism, Or Something Near Enough* O'Connor, T., and Wong, Hong Yu, forthcoming, 'The Metaphysics of Emergence', *Noûs*

Papineau, David (2001) "The Rise of Physicalism" in *Physicalism and its Discontents* ed. Gillett, C., and Loewer, B. Cambridge University Press.

Putnam, Hilary (1975) *Philosophical Papers Vo. 2* Cambridge: Cambridge University Press

Russell, Bertrand (1913). "On the Notion of Cause." *Proceedings of the Aristotelian Society* 13: 1-26.

Ryle, Gilbert (1949) *The Concept of Mind*

Shoemaker, Sydney (2001) "Realization and Mental Causation" in Gillett and Loewer (2001)

Tooley, Michael (1987) *Causation: A Realist Approach*. Oxford: Clarendon Press.

Daniel R. Wegner, D.M. (2002) *The Illusion of Conscious Will* Cambridge: MIT Press, 2002

