

# A Structure for Mental Causation\*

May 29, 2012

## Abstract

The space of solutions to the mental causation problem (as it is articulated by Jaegwon Kim) appears to divide in two: either mental properties are identical with physical properties, or else they are distinct. Identity solutions are difficult to distinguish from eliminativism. Distinctness solutions entail property overdetermination. Kim argues that since physical properties have the stronger claim to being what really makes things happen, overdetermination entails property epiphenomenalism. But there is a third option to consider: mental properties are neither identical with nor distinct from physical properties. This paper investigates this structure—property overlap—in two quite different ontological settings. The aim of the paper is to recommend the overlap structure as an element in a solution to the mental causation problem, whatever your overall ontological preferences may be.

## 1 Introduction

Mental causation (causal relations between mental states and other things, mental and non-mental) has always been philosophically troubling. Dualists like Plato or Descartes provide a concept of mind (or soul) that ensures its ontological specialness; but notoriously it seems impossible for a nonphysical mind to interact causally with the physical world.

Contemporary scientific monism provides compelling evidence that minds cannot make a difference in the physical world. There are two kinds of argument. One is empirical. Huxley (1874) and more recently Wegner (2002) argue that there are detailed physical mechanisms—that is to say: non-mental mechanisms—that fully explain everything that happens involving a mind. The causation of behavior has a common cause structure: body and brain events cause motor events and cause mental events. The mental events do not cause the motor events (or anything

---

\*Thanks to audiences at the Metaphysics of Science conference, Melbourne, AU, July 2009, at the SoPhA conference in Geneva, CH, September 2009, and at the University of Georgia, October, 2009, and two anonymous referees for helpful comments.

else). The best evidence for a common cause structure like this is separability: the brain mechanism can produce the motor event without the mental event (Huxley's Sergeant F——), or the mental event without the motor event (cases of the illusion of control).

The other kind of argument is metaphysical rather than empirical. It appears to be more fundamental than the empirical argument. It is a very old argument: its outlines appear for instance in Plato's *Phaedo* (98c-99c) and Lucretius' *De Rerum Natura*. It is neither about specific mechanisms nor about psychology in particular. It applies to any situation where there are both physical and higher level candidates for causation. Consider Socrates's reasoning in his cell. He thought the better course of action was to remain in Athens. His remaining there has a physical, bodily cause: his sinews and bones apparently cause him to remain. (Socrates, of course, disagrees.) There are two causal explanations: the higher level one refers to what Socrates thinks, and the lower level one refers to the entire physical microstructure of the situation. There are, apparently, two candidate causes for Socrates' remaining where he is. The physical microstructural cause is the better or stronger candidate. The conclusion is then that the physical microstructure is the cause, and the higher level aspect is epiphenomenal: Socrates' belief as to the best course causes nothing. As it is sometimes expressed, the "causal work" is being done by the physical aspects of the situation, not the higher level aspects.

Both of these arguments assume that the mind and body are two, and show that only the body is involved in causation. If we endorse an even more thorough-going monism, and assume that the mind and body are really one, then we seem to lose the mind: there's nothing distinctive about it, not even distinctively mental properties of physical bodies.

This paper will argue that there is a middle ground between dualism and monism. One might call it "one-and-a-half-ism"; it has been termed "new compatibilism" in the literature (Horgan, 1997; Bennett, 2003; Harbecke, 2008). The middle ground is *property overlap*, defined this way: two properties overlap iff they are neither identical with one another nor fully distinct.

The plan of the paper is as follows. First I'll lay out the background assumptions for the metaphysical problem of mental causation, and then briefly review Jaegwon Kim's canonical statement of the problem. I'll develop the idea of the overlap structure using Kim's argument as a foil. To use the structure, we need to take a stance on a network of fundamental metaphysical issues, ones about minds, causation, physicalism, laws, and properties: that is to say, we need

to implement the structure with particular accounts of how causation and so forth work. I'll develop two implementations of the structure in some detail. Each implementation generates a candidate solution to the mental causation problem. Each implementation/solution pair has difficulties that will be familiar to their adherents and their critics. I'll end the paper by commenting on the ontology, and the metaontology, of the overlap relation itself.

The goal of the paper is to offer the overlap structure as a useful tool for exploring responses to the mental causation problem. The goal of the paper is *not* to argue for either of the implementation/solution pairs. I have argued elsewhere (Dardis, 2008) for a solution to the mental causation problem based on the structure. My concern here is not to try to solve the mental causation problem but to recommend the overlap structure: try it out and see if it helps your concerns about mental causation.

## 2 Background

The metaphysical problem of mental causation, as it is currently discussed, depends on some assumptions about our overall metaphysical setup.

- **Physicalism:** every actual concrete individual is made up of nothing but physical matter. “Physical” means: referred to by an ideal completed theory of what makes things happen in the natural world.
- **Closure:** physics is **closed** in the sense that for anything that happens, it has a purely physical cause, if it has a cause at all.
- **Causal relevance:** when one event causes another, the cause has properties *in virtue of which* it causes the effect (alternative expression: one event causes another *because it has* certain of its properties rather than others). Events are “coarse-grained”, that is, they are particulars that have many properties and relations. Given an effect, some properties of the cause matter to the existence and character of the effect, others do not. Call this relation on properties **causal relevance**.<sup>1</sup>

---

<sup>1</sup>The current canonical formulations of the mental causation problem tend to be expressed using a “fine-grained” conception of events, on which two events are identical iff they share the same constitutive object, time and property. There is some discussion in the literature about whether the mental causation problem depends in some way on which conception of events we adopt. I don't believe that it does, so I will assume that the problem can be expressed using

- **Supervenience:** mental properties supervene on physical properties. This means: two things can't differ mentally without also differing physically. I'll follow Kim in using a slightly stronger conception of supervenience: mental properties supervene on physical properties in the sense that, necessarily, any event that has a given mental property has one from a definite range of physical properties, and, if anything has one from that range of physical properties, it has the mental property as well.<sup>2</sup>

With these assumptions, we can formulate a version of Kim's "supervenience" or "exclusion" argument (Kim, 2005). Assume (for *reductio*) that a mental event causes some effect in virtue of having some mental property. And suppose that  $M \neq P$ : the mental property is not identical with the physical property. By Supervenience, the cause and the effect have physical properties that necessitate their mental properties. By Closure, the cause produces the effect in virtue of its physical properties. Hence there are two candidates for which property instance causes the effect: the cause having its mental property, or the cause having its physical property. The Exclusion Principle says that no more than one independent candidate can make the effect happen (except in cases of genuine overdetermination). Since by Closure we already know that the physical cause makes the effect happen, the mental cause is excluded. So the mental event causes the effect only in virtue of having its physical property.<sup>3</sup>

### 3 Dualism, property identity, monism

The ontology sketched in the last section makes room for two kinds of concrete individuals: objects and events. The assumption of **Physicalism** requires that these concrete objects and

---

either conception of events, and use the "coarse-grained" conception.

<sup>2</sup>Some philosophers argue that supervenience needs to be robust and explanatory (Horgan, 1993). Others argue that it cannot bear this explanatory burden (Heil, 1998). The arguments of this paper do not require supervenience to do any explanatory work. I assume supervenience because the conjunction of physicalism and the falsity of supervenience entail that there is no mental causation. The second conjunct of the definition of supervenience is what matters: if it is not necessary that something that has one of the base physical properties also has the mental property, then the physical effect can happen whether or not the mental property is instantiated. How to show that supervenience holds is a task outside the scope of this paper. The assumption of supervenience includes the idea that there are mental properties and that they stand in a certain relation to physical properties; without this idea, the mental causation problem cannot get started (Glymour, 1999).

<sup>3</sup>There is a fairly enormous literature on Kim's argument. My focus here will be on the the sense in which the physical cause and the mental cause are "two". For present purposes I am accepting the rest of Kim's argument without question.

events are made up of nothing but physical matter. On reasonable assumptions **Physicalism** rules out substance dualism, the view that some concrete individuals are not physical. Again on reasonable assumptions **Physicalism** entails at least token-token identity: every mental event is identical with some physical event.<sup>4</sup>

Type-type identity adds property identity to mental-token identity: every mental property is identical with some physical property. A fairly popular strategy for dealing with the mental causation problem recently is to say that mental properties are identical to physical properties. Kim (2005) uses the mental causation problem to argue against non-reductivism and for identifying mental events with physical events (at least for the aspects of the mind that don't involve consciousness and qualia) by way of functional reduction. Heil (2003) holds that there are dispositions and their existence and manifestation constitute truthmakers for propositions about various kinds of things, among them the mental ones. Heil (2009) suggests that this kind of view fits nicely with Davidson's anomalous monism (1970): it's not about *properties*, it's about which *descriptions* go with which laws. Reality doesn't come in layers or levels; only our theories do that.

There is at least a whiff of eliminativism about the property identity view.<sup>5</sup> The view says there are no distinctively mental properties. It's not literally true of you that you have the property of thinking about Australia. But then, if nothing, really, has any mental properties, then there are no minds.

This is not a decisive consideration. According to Heil, it *is* literally true of you that you think about Australia (and hence it is literally true of you that you have a mind). There are dispositions that make this claim true. They have physical descriptions as well.

Yet Heil's view is committed to a distinction between ways that the world makes sentences true. Either (a) objects in the world and their dispositions are the referents of the referring terms and predicates of the sentences; or (b) those sentences are made true in some other way. Truth and existence look second-class on the second alternative: perhaps mental sentences are made true in some way similar to the way sentences about the average family are made true. There's still the whiff of eliminativism. I would prefer a robustly non-eliminative position, and so I think

---

<sup>4</sup>There is room to disagree with this entailment claim, but that won't matter for the arguments of this paper.

<sup>5</sup>Davidson's view—despite being a version of a token identity theory—is not infrequently taken to be a kind of irrealism about the mental. One way to express this charge would be to say that the only “real” level is the one described by the fundamental laws.

we need to say that mental properties are not identical to physical properties.

## 4 Non-identity

The non-identity of mental properties and physical properties is supposed to follow from non-reductivism. The main arguments for non-reductivism are the multiple realizability argument, Davidson's anomalism argument about the rationality of the propositional attitudes, and considerations about consciousness. All of these arguments are controversial and troubled.

What matters for Kim's exclusion argument is the premise that the mental event is not identical with the physical event (" $M \neq P$ "). Depending on how it is understood, reductivism may well be independent of identity. On Ernest Nagel's classic account, reductivism is consistent with non-identity. If reductivism is a view about theory relations rather than ontology, then non-reductivism is consistent with identity.

Here is a neoCartesian modal argument for non-identity. Suppose that non-physical bearers of mental properties are conceivable; and suppose this is enough to show that they are logically possible. Now suppose that properties are individuated with respect to their logically possible bearers, so that if some possible individual  $a$  has property  $P_1$  and lacks  $P_2$ , then  $P_1$  is not identical to  $P_2$ . Then, since we are assuming that some possible individual has mental properties and lacks any physical properties, mental properties are not identical with physical properties.

This argument is consistent with physicalism as I describe it above. It's also consistent with the Jackson formulation of physicalism: any minimal duplicate of our world is a duplicate *simpliciter* (Jackson, 1998, 12). (Hence no *actual* individual who has mental properties is non-physical; nor is any nomologically possible individual with mental properties non-physical.) So a mental property can be non-physical, in this sense, without being in any way "spooky" and without any of the mental individuals in our world (or any world of which our physics is true) having any non-physical parts or accompaniments. Its non-physicality is conceptual, not ontological.

I don't claim this argument is decisive. For the purposes of this paper what matters is that there is some reasonable argument for the non-identity of mental and physical properties. If there is no such argument, then there is no mental causation problem.

**Closure**, as I defined it above, requires that if an event has a cause, it has a physical cause.

**Causal relevance** entails that the physical cause produces the effect in virtue of its physical properties. **Closure** does not rule out that this event also has mental properties; and, most importantly, it does not rule out the possibility that this event causes the effect in virtue of its mental properties as well as its physical properties. (Clearly, if it did rule this out, then the property identity solution to the mental causation problem would be the only possible contender.)

## 5 Non-identity and distinctness

The foregoing highlights the following position in the space of views about minds, bodies and causation: physicalism is true, and so every mental individual is a physical individual. Causation requires the causal relevance of properties. Physical properties are (we are assuming) causally relevant. If mental properties are identical with physical properties, then there is no mental causation problem, but, equally, there is no mental either. But if they are distinct from physical properties, then Kim's exclusion argument seems to go through: there are two candidates for the causation, for example, of an action, and the physical property has the better credentials for counting as the cause.<sup>6</sup> The arguments I've offered to indicate this position do not, obviously, force us to take up this position. But the position is a familiar one, and one that many philosophers have struggled hard to escape from. For the rest of the paper, then, I'll take the following dilemma to set the fundamental problem of mental causation. Given the assumptions articulated above, if mental properties are identical with physical properties, then we have an unacceptable eliminativism; but if mental properties are distinct from physical properties, then we have an unacceptable property epiphenomenalism.

Most<sup>7</sup> arguments to the conclusion that there is no mental causation depend at least implicitly on an exclusion principle, to the effect that in some important way a cause is "one". Kim has made this principle explicit in a series of articles. In recent formulations, Kim's exclusion principle rules out "more than one sufficient cause" (Kim, 2005, 42); in his original formulation it rules out more than one "independent" cause (Kim, 1989, 239). As Bennett (2003) and others

---

<sup>6</sup>Corbí and Prades (1999), Menzies (2009) and others argue that the mental cause excludes the physical cause: in these cases of mental causation the physical isn't a cause and the mental is. I can't argue for this here, but I believe we should hang on to the idea that physical causation is involved in all causation.

<sup>7</sup>Exceptions include Kantian and Wittgensteinian and Rylean arguments to the effect that minds are categorically not in the causing line of business.

have pointed out, the assumption of supervenience entails that the mental cause and the physical cause are not independent. The more recent formulation raises a further question: what if the mental cause isn't strictly distinct from the physical cause—what if, in some useful sense, they aren't strictly speaking more than one thing?<sup>8</sup> If mental properties are neither identical to physical properties nor fully distinct from them, then the Exclusion Principle doesn't bear on them, and so Kim's argument doesn't go through.

What on earth could it mean to say that property  $P$  is neither identical with property  $Q$  nor not identical with it?

When things can overlap, there are three possibilities, given individuals  $a$  and  $b$ :  $a = b$ ;  $a \neq b$  and  $a$  and  $b$  are entirely distinct (share no parts); or  $a \neq b$  and  $a$  and  $b$  share a part. In the last case, where  $a$  and  $b$  overlap,  $a \neq b$  (they are not identical) but neither are they fully distinct. (See Sanford (2005) for a nice discussion of the distinction between non-identity and distinctness.)

## 6 The overlap structure

Let us then define “property overlap” as follows: two properties  $P$  and  $Q$  overlap if and only if  $P \neq Q$  and  $P$  is not fully distinct from  $Q$ . (What exactly it means to say that one property is not fully distinct from another depends on what properties are, and is hence an aspect of what I'm calling an “implementation” of this structure, not something to specify in advance of an implementation.) My suggestion for a structure for mental causation is this: mental properties overlap physical properties.

Some comments are in order before we show how to use the structure.

By calling it a “structure” I mean to abstract away from almost all details about what overlap actually is. The reason for this is that my aim in this paper is to show how the structure can be

---

<sup>8</sup>Kim derives the claim that “ $M \neq P$ ” from irreducibility. Kim does not discuss what exactly this claim means (other than in the mysterious footnote 9 at (Kim, 2005, 42), which appears to conflict with one of the central tenets of non-reductivism, token-token identity; see (Fiorese, 2009)). It could mean 3 that the instance of property  $M$  is numerically distinct from the instance of  $P$ ; or it could mean that they are not strictly identical but they are not numerically distinct either. I am in effect faulting Kim's argument for overlooking this distinction, and inferring numerical distinctness from non-identity. But suppose Kim *isn't* deriving numerical distinctness from non-reductivism, and suppose he thinks that the two property instances are not identical yet not distinct. It then becomes puzzling how the various formulations of his Exclusion Principle could be brought to bear in the argument: (1) do we have “more than one” sufficient cause? (2) are they independent? I think the more charitable reading has Kim making a natural-seeming inference, from non-reductivism to non-identity to numerical distinctness; a less charitable reading has him *not* making this inference yet remaining oblivious to the consequences for his argument. Thanks to Peter Menzies for raising this question.

filled in or implemented in a variety of ways.

The overlap structure is a structure for ontology. By itself it is not a solution to the mental causation problem. To get a solution, or a class of solutions, we need also an account of causation and how it is related to these overlapping properties, that shows how an instance of an overlapping pair of properties can make things happen in virtue of both of them. (It is, of course, possible to endorse overlap for properties but deny that it has any useful connection with causation: for example, by saying that all causation is defined over properties that do not overlap.)

I think that if the structure can be successfully implemented, the resulting package should have clear advantages over some other accounts of mental causation. It is not eliminativist. The overlap structure permits mental properties to be non-identical with physical properties, and hence permits retaining a kind of ontological distinctiveness for the mental. The overlap structure is intended to be fully compatible with physicalism. It is consistent with the claim that all mental events are physical events: that all mental events have physical properties. What these events do is entirely accounted for by (ideal, completed) physics. The overlap structure thus offers a considerable advantage over dualisms of various sorts, including ones that otherwise honor my assumptions, as does pre-established harmony.

The overlap structure is intended to be resolutely and robustly metaphysical (as opposed to epistemological): whether two properties overlap is a fully objective, mind-independent fact about them.

The overlap structure is compatible with property overdetermination: a given effect of a mental event has (at least) “two” causes, one physical, one mental (scare-quotes around “two” since they are not actually fully distinct). The relevant sort of overdetermination would be property overdetermination, defined as follows: properties  $P$  and  $Q$  of event  $e$  are overdetermining causally relevant properties with respect to effect event  $f$  iff  $P$  and  $Q$  are both causally relevant to some property  $R$  of  $f$ , and  $e$  would have caused  $f$  if it had had  $P$  but not  $Q$ , and if it had had  $Q$  but not  $P$ .

But the overlap structure does not entail property overdetermination. If we assume **Supervenience**, for example—and I’ve argued above that we must, if there is any hope for a solution to the metaphysical mental causation problem—it is not nomologically possible for  $e$  to have  $P$

but not have the mental property *M* that supervenes on it.

It may be objected that only by a technicality do the combination of overlap and **Supervenience** not entail overdetermination. This is a powerful objection. It is at least implicit in much of what Jaegwon Kim has written about mental causation and exclusion. We can express it as follows. Suppose *M* is not distinct from *P*, but overlaps it. And suppose it is conceded that *P* causes the effect. *M*, in this picture, isn't doing anything. The most economical account of how the world works refers to the physical properties. It has no need to refer to the mental properties.

I think Kim is right that a certain kind of theoretical economy exerts a very strong intellectual pressure on us to claim that the working of the world is purely physical. Philosophers can and do dispute the exact nature, and the acceptability, of this idea of economy. The pressure seems to remain.

However, I do not believe that this idea of theoretical economy is itself logically necessary, or obligatory for any correct metaphysical theorizing. There are alternative ways to set up a metaphysical theory of properties, laws and causes, on which not all causation is "fundamental". I think the strongest reason for seeking such alternatives is that they can accept a great deal of "common sense" about causation: dropping a wine glass on a granite counter causes it to shatter, days of drought cause the plants to wither—and choosing to do the right thing makes one try to do the right thing. (Again, such a metaphysical theory only removes one metaphysical difficulty. It could, unfortunately, turn out that there is empirical reason to believe that some or all common sense causes are really instances of a common cause structure, and hence are really only "images of causation".)

One of the main obstacles in the way of these less economical alternative metaphysical theories is the idea that fundamental properties are fully distinct from higher level properties. If they are, then there are genuinely two candidates for what does the causal work, and it will then seem obligatory to pick one of them. If they are strictly identical, then we lose the idea that the world really contains anything other than the fundamental entities. I recommend the middle way: mental properties are not strictly identical with physical properties, but neither are there genuinely two candidates: so we do not have to pick one of them as really doing the causal work.

## 7 Candidate solution #1: properties as classes

What I am calling an “implementation” of the overlap structure is a specification of the natures of properties and causation (and perhaps other features of the metaphysical landscape) that makes room for property overlap. Calling the overlap structure a structure for mental causation indicates that an implementation should have consequences, preferably happy ones, for the mental causation problem.

In this section I’ll sketch a neoHumean implementation of the structure, and hence a neo-Humean candidate solution to the mental causation problem. My aim, again, is not to argue for this solution. Rather, my aim is to provide an example of one way to implement the structure for mental causation that I am offering. The abstract overlap structure is not committed to any particular such implementation. In Section 8 I’ll show how to implement it using dispositions and the process account of causation as the ontology of properties and causation, and Section 9 gives a brief catalog of some other overlap solutions that have appeared in the recent literature.<sup>9</sup>

Assume, then, a neoHumean ontology rather like David Lewis’s (with some significant modifications). Space-time points are the fundamental individuals. Other individuals (like me) are mereological sums of fundamental individuals; many (again like me) are causally unified in various ways. Properties are sets of actual and possible individuals. Set-like operations on properties are defined: for example, the disjunction of two properties is their union. There’s nothing “disjunctive” about that set, though: it’s just another set of individuals. Similarly functional-role properties: the property of having a property that can cause shattering is a collection of individuals, not a “second order” property.

The laws of nature are a set of regularities over these properties: the strongest simplest set of universal generalizations from which all the particular matters of fact may be derived. This is the Mill/Ramsey/Lewis (MRL) “best system” account of laws. Lewis (1983) argues for an additional constraint: the properties over which the laws are defined are to be natural properties, in Lewis’s proprietary sense of naturalness. Call “nomic relevance” the relation on any set of properties

---

<sup>9</sup>In (Dardis, 2008) I argued in some detail for essentially the solution to the mental causation problem I will describe in this section of the paper. The ontology is neoHumean, and hence problematic for the many antiHumean metaphysicians working today. But the structure is not essentially bound up with the ontology, and, as will emerge, can be given an antiHumean implementation as well. Hence my goal in this paper: to suggest that the structure is interesting and valuable, even if my favored use of it, in a neoHumean framework, is hampered by that neoHumeanism.

related by a law of nature.

Events are changes in properties. If one event causes another event, then the two events have physical properties such that those properties stand in the relation of nomic relevance (Davidson, 1967, 1995). To a first approximation, causal relevance, as defined above, is nomic relevance, so the fundamental nomic properties are causally relevant. So far, though, causal relevance is exclusively fundamental: it only relates properties mentioned by the laws of nature, which are—so far—fundamental laws.

Naturalness comes in grades. Being a visual experience of an apple is less natural than being a molecule of 11-cis retinal, but far more natural than being grue. Regularities can be formulated over less than fully natural properties. (Indeed, Lewis's argument for naturalness depends on a non-uniqueness claim about laws and properties: if there is one set of regularities that forms a best system, there are infinitely many. There are fewer such regularities referring only to natural properties.)

The standard picture of how “macro” properties are related to fundamental physical properties is that “higher level” properties are multiply realized by physical properties. For example, the property of being a cat can be realized by different total configurations of physical matter: this one is a tabby, that one a calico, this other one has lost its ear, and so forth. Hence instances of macro, less than fully natural, properties will differ physically, and hence differ in what exactly they can do. So laws of nature that refer to mental properties will need to be “qualified”: the mental property is a partial specification of a complete physical way something can be, and the qualification adds in the rest.<sup>10</sup>

The obvious way to do this qualification is simply to specify the complete physical characterization of the instances of the mental properties. But that would be to just give up on the idea of laws of nature that refer to mental properties, since the mental properties drop out of the complete physical characterization.

Let us think instead of the mental property as something like an Aristotelian “form”: an incomplete specification of a way things might be. For any instance of the form, there is a cor-

---

<sup>10</sup>Functionalism is one model of how a macro property can be multiply realized. The essence of a functional property is causal: for example, in the toy example, the property of being pain is the property of being caused by bodily damage and causing avoidance behavior. But there are other ways for a property to be multiply realized, for example for its essence to be relational in some other way than causal.

responding “matter” property, which collects all individuals that have the physical stuff that this individual has, in whatever form that physical stuff may occur. This “matter” property isn’t physical in the sense defined in Section 2, since no (natural) ideal completed theory of everything is going to refer to these matter properties. But the conjunction of this form property and this matter property determines a physical property that is referred to by the laws of nature. We can then express the laws of nature as regularities over higher level properties with qualifications like these added in. Higher level laws have the disadvantage that they are not so simple as the fundamental laws, and the advantage that they are stronger, in that they display higher level regularities.

We then add properties related by higher level sets of laws to the causal relevance relation. These higher level sets of laws are built up around nearly natural properties. “Nearly” here is be connected with “suitability for serious science,” so that a higher level biochemical type like 11-cis retinal is nearly natural, a merely instrumental higher level psychological type like intelligence may not be natural enough, and (if there is mental causation), a higher level psychophysical type like willing to move one’s hand is natural enough.

## 7.1 Discussion

A few comments on this first candidate implementation are in order.

First, property overlap is well-defined. Properties are sets of actual and possible individuals. Since we are assuming supervenience, every nomologically possible mental individual is some physical individual, and so the mental property is a proper superset of the physical property. This fact enables the candidate solution to respond to Kim’s exclusion argument: since the mental event and the physical event do not constitute two independent causes of the effect, the exclusion principle does not apply to them. Hence the candidate solution is an example of what I am calling the “overlap” structure for mental causation.

Second, the candidate solution adopts a proposal about what causal relevance is, and shows how it could be satisfied in the case of mental causation. Causal relevance is nomic relevance: if one event causes another, it does so in virtue of some of its properties rather than others, and the ones that matter are the ones that are linked to effect properties by laws of nature. The candidate solution avoids “fundamentalism” about causal relevance by assuming that the causal

relevance relation includes properties related by laws of nature involving nearly natural properties.

Third, this is a neoHumean account of properties, laws and causation, and neoHumean accounts of these things have been sharply criticized in recent years (Armstrong, 1983; Ellis, 2001). This account holds that properties are sets: in this picture, even being an instance of an intrinsic property somehow involves other instances of the property. Its conception of properties is naturally taken to be categoricalist, and to construe properties as quiddities (but see below, Section 8.3), since the laws of nature are, at least according to Hume, contingent, permitting apparently objectionable “transubstantiation”: there are possible worlds where the property of mass and the property of being beer exchange their nomological roles. This account holds that laws are only regularities over what in fact happens: laws cannot in this picture explain why one thing makes another thing happen. (One popular expression of this concern is that Humean laws have no “oomph”.)

The overlap structure can be implemented with an ontology of laws and properties that does not have these objectionable features. In the next section I’ll sketch how it would look with an ontology of dispositions. Again, my aim is not to argue that this implementation is preferable, or that the overall package is a better solution to the mental causation problem. My aim is rather to explore (and to offer and to recommend) the overlap structure as a solution to the problem of mental causation.

## **8 Candidate solution #2: Properties as dispositions**

The overlap idea looks particularly ill-suited for ontologies where there are no levels—in which there is some basic category of “truth makers” such that it doesn’t make sense to talk of the mental and the physical as being at different levels, and consequently it doesn’t make sense to talk of the mental as being not identical identical with (either as fully distinct from or overlapping the physical). Heil (2005) offers a framework for a “no level” conception of dispositions. In this section I will investigate how his framework might make use of the overlap structure.

## 8.1 What are dispositions?

Dispositions are actual intrinsic features of things. Subjunctive conditionals may tell us something about a disposition, but the disposition is something distinct from those sentences—its nature is not exhausted by some set of subjunctive conditionals. A disposition fully fixes what its manifestations can be like; in other words, the relation between a disposition and its manifestations is not contingent. In particular, it is not contingent on what laws of nature happen to hold. Rather, law sentences are made true by the dispositions. Since the relation between disposition and manifestation is necessary, the laws of nature are necessary.<sup>11</sup> Event causation is a matter of the manifestation of dispositions.

According to Heil, dispositions are not grounded in something “lower level.” There cannot be levels of dispositions. The argument is closely related to Kim’s exclusion argument (347-350). Assume for *reductio* that there can be higher and lower level dispositions. Heil sees a trilemma: either the effect is overdetermined by both the more and the less fundamental disposition; or the lower level disposition is not by itself sufficient for the effect; or else one of the purported causes really isn’t a cause. The first is implausible. The second violates closure. On the third alternative, since the more fundamental disposition is the better candidate for making the effect happen, the higher level disposition would be powerless. Since Heil assumes that dispositions are always “powerful”, it then follows that there are no higher-level dispositions.

## 8.2 Using the overlap structure

The argument depends on the assumption that the more fundamental and the less fundamental properties are distinct (cf. the expression (349) “higher-level items”). Can we give this up? Heil comments at the start of his piece (343) that “‘Disposition’ is a term of art: you can define dispositions as you please”. So we have some liberty in setting up an account of dispositions that we can use in a “unified understanding of mind and world” (351).

Let us then say that there are physical dispositions and also mental dispositions (and others, of course). The mental ones overlap the physical ones: they are neither identical with nor fully distinct from the physical dispositions. We don’t have levels, since we don’t have any dispositions

---

<sup>11</sup>The possibility of interference makes the relevant notion of necessity problematic; see (Schrenk, 2011).

(fully) distinct from the physical dispositions.

How should a dispositionalist explain what it is for two dispositions to be not fully distinct, yet not identical either? There seem to be only two avenues. One is to say that dispositions are theoretically fundamental, and that their overlap is similarly fundamental: the overall metaphysical theory takes it as axiomatic that properties can overlap, and there is no more to say. The other is to say that dispositions have some kind of theoretical structure. Consider Sydney Shoemaker's account of properties (Shoemaker, 1984): they are sets of conditional causal powers. Suppose dispositions are essentially associated with a set of causal powers, and are individuated with respect to that set. Then  $d_1$  overlaps  $d_2$  when the intersection of the associated sets of causal powers is not empty, but is not identical with their union.

### 8.3 What “does causal work”?

So far, all we have is an answer to the exclusion argument, and an answer to the argument that there are no levels: since the mental dispositions and the physical dispositions aren't distinct, they aren't independent causes, and so the exclusion argument doesn't apply to them (the third horn of the trilemma is blunted). But we haven't explained how both can be “powerful”: we haven't begun to explain what it might mean to say that a given event is caused by (is the manifestation of) both a physical disposition and a mental disposition.

Again, the overlap structure by itself doesn't answer the question how mental dispositions “do their work” or make things happen. The exclusion argument was supposed to show that the supervening candidate *could* not cause the effect, since there can't be more than one non-overdetermining independent cause. Assuming that the mental dispositions overlap the physical ones, there are two candidates for causing the effect, which are not independent, and not overdetermining. It clearly doesn't yet follow that they *do* both cause the effect. In fact, there appears to be a perfectly good reason to say that they *don't*: the physical cause is sufficient, if anything is, to produce the effect. The mental candidate would be “‘pretend’ or ‘faux’” (Kim, 2005, 62), nothing new or different from the physical cause.

The point is exactly parallel to one that came up in the Lewisian implementation in Section 7 above. According to Lewis's formulation of the MRL view, a law of nature is a sentence from a strongest simplest theory of everything. We can define the notion of “causal work” or “mak-

ing something happen” in terms of the laws of nature: the causally relevant properties are the nomically relevant properties from the fundamental laws, and that’s it. Whether or not mental properties overlap physical properties, clearly the only properties that contribute to “causal work” according to this definition are ones referred to by the laws of nature, which will be the laws of fundamental physics. (Thus a broadly Humean picture of what “making happen” means can be just as “fundamentalist” as any other picture, and hence just as exposed to the mental causation problem.)

## 8.4 What does “doing the causal work” mean?

This revived exclusion-style argument assumes something critical without explicating what it is. Physical causes are the best candidates for being causes, and when they compete against higher level causes, they always win. But what is it about physical causes that explains how they make things happen—what does “doing the causal work” mean?

### 8.4.1 Humean answers

The textbook account of Hume’s answer to this question is, of course, “nothing at all.” The only thing it could mean, according to Hume, is “necessary connexion”, but, it turns out, those words don’t actually correspond to any idea (or: they don’t mean anything).

Humeans other than Hume answer the question with fancy versions of “constant conjunction”. They can, as we just saw, be fundamentalists. They don’t have to be. As we saw above in Section 7, there is a fairly straightforward way to define a concept of natural law, and causal relevance, that makes room for causal relevance at many levels.

### 8.4.2 Non-Humean answers

As usual, anti-Humeans will reject both the fundamentalist and this more ecumenical account of causation as not seriously providing any answer to the question of what causal work is. We want “necessary connexion”; we want to know how the cause *makes* the effect happen.

One way to respond to this demand is to show that the relation between causes and effects is necessary. There is a variety of proposals that make this demonstration: Sellars (1948) argues in effect that properties are individuated with respect to laws, and hence what they do is essential

to them; Shoemaker (1984) argues that causal powers necessitate their effects; Armstrong (1983) argues that a law expresses a necessitation relation between universals (although he also argues that whether this relation holds is itself contingent); Bird (2001) offers the “down-and-up” argument to show that salt’s dissolving involves exactly the same physical mechanism as water’s being able to dissolve, and hence it’s not possible for *salt* not to dissolve in *water*. Ellis (2001) argues, as does Heil, that what dispositions do is essential to them.

### 8.4.3 Necessity isn’t “connexion”

Suppose that these arguments do establish that the connection between an instance of a property or a disposition and an effect is a necessary one. Then Hume was wrong that there is nothing necessary here. But do we now know what the *connexion* is? I don’t think we do. Take Bird’s lovely demonstration. Salt (sodium chloride) is a crystalline molecule held together by ionic bonding. The molecules in the lattice will experience electrostatic attraction and repulsion. Hence Coulomb’s law is true of salt. Similarly, Coulomb’s law is true of water. Hence the law that requires salt to dissolve in water follows from what salt and water *are*. Hence dissolving is necessary. Now: why do atoms experience electrostatic attraction and repulsion? The answer takes us down a level, to electrons and their charge, and the geometry of the molecules. How does charge *work*, i.e., how does electrostatic attraction bring together a negative and a positive ion? Well, there’s more to say about the structure of what charge does. But, I think, in the end, there is an end to what we can say. At that end, the answer to, “how does that *work*?” is, “it does”.<sup>12</sup>

Notice also that a purely Humean account of causation and laws has the resources to hold that the connection between cause and effect is necessary. If properties are individuated with respect to the laws that govern their instances, then there is room for a position on which property instances necessitate their effects. Take Lewis’s conception of properties as sets of actual and possible individuals, and take some law of nature, for instance Coulomb’s law. If the properties referred to by Coulomb’s Law are sets the members of which are exactly the actual and possible individuals that satisfy the law, then the law turns out to be necessary and the properties necessitate their effects. (Even better, these properties are in no sense quiddities.) This explanation

---

<sup>12</sup>Unless all of physics is grounded in some deeper necessity (as Einstein hoped that it all reduces to geometry).

of why the connection is necessary leaves what I take to be the disturbing conclusion of Hume's critique of causation untouched: there is no *connexion* between the cause and effect, even though the effect necessarily happens, given the cause.

#### 8.4.4 Overlap, dispositions, and process accounts of causation

Suppose there is some account of "making happen" that is defined over properties or commonalities among events (rather than a singularist account). I know of roughly three such accounts in the recent literature: (1) counterfactual accounts (Lewis, 2000); (2) mechanism accounts (Glenan, 1996); and (3) process accounts (Dowe, 2000). I myself find counterfactual accounts unpromising, since it seems as though the truth of the relevant counterfactuals has to depend on something else, like laws, or dispositions. Similarly mechanisms seem to me to be explained on the basis of laws or dispositions, and not the other way around. So for the purposes of this investigation of how to implement the overlap structure with dispositions, let me consider what it would be like to put a process account of causation together with a view that holds that dispositions may overlap.

The process account of causation holds that when one event causes another, there is a causal process that connects them. A causal process transmits a conserved quantity between the cause and the effect. Conserved quantities are defined in terms of physical properties, like mass-energy, momentum and charge (Dowe, 2008, Section 5).

Putting this together with the basic dispositional account, we get, one event causes another when the first has a disposition, and the manifestation of that disposition involves the transmission of a conserved quantity. So take a case of mental causation: the desire for the cheesecake causes Suzy to reach for the cheesecake. There's some collection of physical dispositions at work here; let's call the collection  $P$ .  $P$ 's causation of the arm to move involves conserved physical quantities.

Now suppose that  $M$  supervenes on  $P$ , and overlaps it. If  $M$  can make  $E$  happen, then there is some quantity conserved when  $E$ 's follow  $M$ 's. Two constraints seem necessary:

- **Non-additivity:** the quantity conserved in the  $M/E$  transaction and the quantity conserved in the  $P/E$  transaction are not independent. In particular, they must not be "additive". If, say, the total energy of the physical basis for the desire is conserved as it moves

the hand toward the cheesecake, *and* some quantity is conserved in the desire/action causal transaction, the quantities don't add up; the effect doesn't end up with *more* of anything than either quantity bestows;

- **Non-triviality**: the existence of this conserved quantity should be non-trivial. The conservation laws in mechanics are non-trivial in the sense that they are a structural element in the theory of mechanics (cf. John Norton's complaint, (Dowe, 2008, Section 6.2)).

About **Non-additivity**, since  $M$  and  $P$  overlap, they aren't distinct; so we could think of the two conserved quantities as stemming from something like two alternative book-keeping schemes for the same transaction. I don't think **Non-additivity** is automatically satisfied given overlap. There are two sciences here; we would have to explain on their basis why the two quantities are not independent. By contrast, *without* overlap, such an explanation would be a great deal harder, since if the two properties are independent, they would appear to make independent contributions to causation. **Non-additivity** is entailed by **Closure**, since the relevant sense of addition would permit a mental cause to make a difference that has no physical cause. But we need more than this for a satisfactory account: we need to explain why exactly **Non-additivity** holds.

If the **Non-triviality** constraint is satisfied, then there will have to be a science of the domain in which  $M$  falls—in this case, psychology must be a science. And this science must describe psychology as involving causation and conserved quantities. Is there any chance whatsoever of psychology being like this? (Dowe, 2008, section 6.6) writes, "in any case, to suppose that the conserved quantity theory will deal with causation in other branches of science also requires commitment to a fairly thorough going reductionism, since clearly there is nothing in economics or psychology that could pass for a conservation law". Dowe's argument here is that it just looks awfully unlikely that economics or psychology have anything like conservation laws. That seems right, but perhaps there is more to be said.

Davidson (in 1985 for one example) took the analogy between propositional attitude psychology and measurement theory very seriously. The argument for the indeterminacy of interpretation he gave is an invariance claim: if there is one theory of the meanings of the words of an agent, and hence one assignment of contents to the agent's propositional attitudes (her beliefs, desires and intentions, and any others if there are any) then there are infinitely many.

Davidson was explicit: this is just like the indeterminacy of temperature, or any other measured quantity.

The analogy is suggestive, but not compulsory: to my knowledge, no one has figured out what the relevant representation and uniqueness theorems would be for propositional attitude psychology.

Suppose, however, that there are dynamic invariances for propositional attitude psychology and decision theory. Now, Emmy Noether demonstrated a deep connection between symmetries and conservation laws: “any differentiable symmetry of the action of a physical system has a corresponding conservation law” (Wikipedia).

The upshot would be—and I acknowledge that this is entirely speculative—that there are conserved quantities for propositional attitude psychology. Their mathematical structure would undoubtedly be very different from more familiar conserved quantities. In particular, the fact that propositional attitude psychology is so “open” (minds are so easily disrupted, thrown off course, by non-psychological causes) might mean that the conservation laws are probabilistic; or perhaps they are less than strict in some other hard to describe way.

## 8.5 Discussion

That completes my investigation of implementing the overlap structure with dispositions and a process account of how dispositions make things happen. The main difficulty that emerged was that the process account of causation doesn’t lend itself naturally to the idea of causation at several levels. That could be built up into a new argument for a one (or no-) level account of dispositions. Interestingly, however, the difficulty is not exclusion anymore: it is rather the process account’s demand that causation involve a non-trivial conserved quantity.

In any event, as I’ve suggested, the process account of causation does not *rule out* causation happening at several levels. There may be a way to modify it away from what Dowe calls “a fairly thorough going reductionism”.

It is important to highlight that the difficulty emerged from the demand for an account of the “making happen” relation. The process account of causation is of course not the only such account. Armstrong (2001) for example argues that we can directly perceive the connexion that Hume said wasn’t there. And a dispositionalist need not accept the demand for an account at all:

she can hold either that “making happen” is a primitive, or that the demand is simply illegitimate.

## 9 Other implementations

Implementing the overlap structure requires specifying two kinds of ontological theory. One is a theory of what might be called “generalities,” things like properties, or dispositions, or aspects, or tropes, or universals. The other is a theory of how those things are related to how the world progresses: that is, to theories of laws and causation. The overlap structure is defined over the theory of generalities two: the generalities must be able to overlap (that is, two can be neither identical nor non-identical). It becomes a structure for mental causation when the theory of causation is added in. If the nature of causation is sensitive to the possibility of generality overlap, and in a way that permits both fundamental and non-fundamental causal relevance, then we have a solution to the mental causation problem.

Humean implementations suffer from the problems of Humean ontologies, primarily the concern that regularities alone do not suffice to explain causation. Anti-Humean implementations are supposed to do better on that score.

Here is a brief tour of some other implementations of the overlap structure.

- Constitution. Pereboom (2001, 2002) suggests that the causal powers of instances of mental properties are constituted by the causal powers of the instances of their realizers.
- Coincidence. Yablo (1992) argues that mental events coincide with physical events—they have all their categorical properties in common, but differ in their essential properties. Causation, then, is governed by a proportionality constraint. So a mental event is suitably proportional to an action, while the physical event on which it supervenes is not. Harbecke (2008) offers a sophisticated development of Yablo’s ideas.
- The “subset” strategy. Watkins (2002), Clapp (2001), and Shoemaker (1998, 2007) suggest that the causal powers of multiply-realized properties are the intersection of the causal powers of their realizers.
- For the purposes of this paper I am assuming that what is important and interesting about the overlap structure is that it provides a clean answer to exclusion worries: if mental

properties (and mental events) are neither fully distinct from nor identical with physical properties (events), then there are not two items that need to be ranked for their contribution to causation. Suppose that supervenience holds as a matter of nomic necessity, so that physical properties do nomologically necessitate mental properties. Then there is no nomologically possible difference experiment that could show the effects of instances of the physical property that are not instances of the mental property. Since difference experiments are one key diagnostic of causal relations, this kind of nomic covariation might be counted as a version of the overlap structure: even if the mental properties are fully distinct from the physical ones, that distinctness is “extra nomological” and so perhaps doesn’t count as relevant to counting causes in the way demanded by exclusion principles. (Shapiro and Sober (2007) argue for a response like this to epiphenomenalism, and Baltimore (2010) criticizes their approach; see also my (Author, Year).)

## 10 Overlap: ontology and metaontology

Put formally, what I’ve proposed is a structure, an abstract relation, on properties: two properties  $P$  and  $Q$  overlap if and only if  $P \neq Q$  and  $P$  is not fully distinct from  $Q$ . I have also suggested that there can be such pairs of properties. I hope to have exhibited the usefulness of the structure in thinking about the mental causation problem.

One of the things left unspecified by the overlap structure is the nature of overlap itself. I have noted in passing that for some implementations of the structure, the nature of overlap is well-defined, and that for others, some choices need to be made.

For the Lewisian account of properties, two properties overlap when they have at least one common member. Overlap is well-defined: it is the same thing as having a non-empty intersection.

The Shoemaker view of properties has a similar structure and virtue. A property bestows a set of conditional causal powers on its instances. So a “part of” relation, or something like it, is defined on properties and causal powers. The “subset” account of higher level causal powers exploits this, and holds that the causal powers of a multiply-realized higher level property is the intersection of the causal powers of its realizers.

David Armstrong entertained a theory of universals on which universals have parts. He used this theory to explain similarity: the reason that red is similar to orange is that the universal for red shares parts with the universal for orange. The two universals are related by “partial identity” or overlap: they share some but not all parts. Lewis pointed out that the way the “parts” of universals compose the universals can’t be the way that parts compose a mereological sum: for example, the universal for hydrogen would have to occur twice in the universal for water. Armstrong’s response was to propose a non-mereological notion of the composition of constituents rather than parts.<sup>13</sup>

Pereboom suggests that the causal powers of instances of higher level properties are constituted by the causal powers of their realizers. The concept of constitution for material objects is controversial and under development; there is rather less discussion of constitution for abstract objects.

Tropes are designed, as it were, for maximal particularity. Properties are intensional or hyper-intensional aspects of particulars, and a trope is a particularized property. The ontological order goes from tropes up to particulars and to properties (two kinds of classes of properties). It might seem natural, given their fundamental role in the trope theory ontology, to think of tropes as unstructured, and hence not the sort of thing that can overlap. But it is not logically obligatory to think of them that way; tropes, like dispositions, are theoretical entities and we have considerable leeway in how we set them up.

We have a rough ordering of accounts of “generalities,” from those on which a very clear overlap relation is defined to ones where it is less clear what it would mean to say that two of them overlap. The farther along in this ordering we go, the more work we need to do to explicate and defend the notion of overlap. For example, we might, for familiar reasons, object to the Lewisian idea that properties *are* sets. One retreat would be to say that there is at least one property for each set of actual and possible individuals, but that the properties are distinct from the sets. Then we could say that those properties overlap when the corresponding sets share members. What exactly does it mean here to say that these properties overlap? It might mean no more than what we have just said: two properties overlap iff the intersection of the sets of their

---

<sup>13</sup>Armstrong’s relation of partial identity on universals is different from the overlap of properties sketched above in Section 7. The overlap of two properties is that set of concrete individuals that have both properties, rather than that set of abstract property components that the two properties have in common.

members is not empty. Or it might mean more; that would depend on what else we say about the nature of these properties.

The general problem here is partially ontological, partly metaontological. How do we go about answering questions about the nature of properties? Do the questions even have answers? For example, if we have a theory of properties like the one just mentioned, must there be more to the nature of property overlap than that at least one individual has both properties?

Once again, strictly speaking, the answers to these questions are independent of the overlap structure itself. They must flow from other ontological assumptions. Suppose someone has a theory of properties on which property overlap is impossible. Then the overlap structure is not useful. Suppose another philosopher has a roughly Quinean attitude toward metaphysical theories (Quine, 1948) and has reasons to allow for property overlap (someone like Lewis himself, for example). Then the overlap structure is useful. Suppose a third philosopher is sympathetic with the general Quinean methodology but thinks that for at least some theories of abstract objects, there really is no truth of the matter about whether the theories are true or false (Balaguer, 1998; Yablo, 1998, 2009). Again, the overlap structure could be useful. And suppose finally that a fourth philosopher is strongly realistic about abstract objects and has a demonstrably productive method for arriving at truths about them; and suppose that that method shows that properties can overlap and that the theories of laws and causation integrate property overlap, either in some way similar to one of the ways I sketched above, or in some other way. Such a theory should also have the resources to give substantive answers to questions about what exactly property overlap consists in.

## 11 Conclusions

If mental properties are distinct from physical properties, the mental causation problem looks hopeless; if they are identical with them, the cost of mental causation is eliminativism. So, I've suggested, it's time to try out saying that mental properties are neither identical with nor distinct from physical properties: they overlap. Overlap by itself does not yield causation or "making happen". More is needed: an account of properties are related to law and causation. This structure can be implemented in a wide variety of metaphysical (and metametaphysical) settings.

## References

- Armstrong, D. M. (1983). *What is a Law of Nature?* Cambridge University Press, Cambridge.
- Armstrong, D. M. (2001). Going through the open door again: Counterfactual vs. singularist theories of causation. In Preyer, G., editor, *Reality and Humean Supervenience: Essays on the Philosophy of David Lewis*, pages 163–176. Rowman and Littlefield, Lanham.
- Balaguer, M. (1998). Attitudes without propositions. *Philosophy and Phenomenological Research*, 58(4):805–826.
- Baltimore, J. (2010). Defending the piggyback principle against Shapiro and Sober’s empirical approach. *Synthese*, 175:151–168.
- Bennett, K. (2003). Why the exclusion problem seems intractable, and how, just maybe, to tract it. *Noûs*, 37(3):471–497.
- Bird, A. (2001). Necessarily, salt dissolves in water. *Analysis*, 61(4):267–74.
- Clapp, L. (2001). Disjunctive properties: Multiple realizations. *The Journal of Philosophy*, 98(3):111–36.
- Corbí, J. E. and Prades, J. L. (1999). *Minds, Causes, and Mechanisms: A Case Against Physicalism*. Blackwell Publishers, Oxford.
- Dardis, A. (2008). *Mental Causation: The Mind-Body Problem*. Columbia University Press, New York.
- Davidson, D. (1967). Causal relations. In Davidson (1980), pages 149–162.
- Davidson, D. (1970). Mental events. In Davidson (1980), pages 240–260.
- Davidson, D. (1980). *Essays on Actions and Events*. Oxford University Press, New York.
- Davidson, D. (1985). A new basis for decision theory. *Theory and Decision*, 18:87–98.
- Davidson, D. (1995). Laws and cause. In *Truth, Language, and History*, pages 201–219. Oxford University Press, Oxford. Originally published in *Dialectica* 49 (1995), pp. 263-279.
- Dowe, P. (2000). *Physical Causation*. Cambridge Studies in Probability, Induction, and Decision Theory. Cambridge University Press, Cambridge, UK.
- Dowe, P. (Fall 2008). Causal processes. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*.
- Ellis, B. (2001). *Scientific Essentialism*. Cambridge Studies in Philosophy. Cambridge University Press, Cambridge.
- Fiorese, R. (2009). Events, functional reduction, and the supervenience argument: Kim not near enough. Presentation for the Metaphysics of Science conference, Melbourne, AU, 2009.
- Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis*, 44:49–71.

- Glymour, C. (1999). A mind is a terrible thing to waste—critical notice: Jaegwon Kim, *Mind in a Physical World*. *Philosophy of Science*, 66(3):455–471.
- Harbecke, J. (2008). *Mental Causation: Investigating the Mind's Powers in a Natural World*. ontos verlag, Frankfurt.
- Heil, J. (1998). Supervenience deconstructed. *European Journal of Philosophy*, 6:145–155.
- Heil, J. (2003). *From an Ontological Point of View*. Oxford University Press, Oxford.
- Heil, J. (2005). Dispositions. *Synthese*, 144:343–356.
- Heil, J. (2009). Anomalous monism. In Dyke, H., editor, *From truth to reality: new essays in logic and metaphysics*, pages 85–98. Routledge.
- Horgan, T. (1993). From supervenience to superdupervenience: Meeting the demands of a material world. *Mind*, 102(408):555–586.
- Horgan, T. (1997). Kim on mental causation and causal exclusion. *Philosophical Perspectives 11: Mind, Causation, and World*, pages 165–184.
- Huxley, T. (1874). On the hypothesis that animals are automata and its history. In *Selected Works of Thomas H. Huxley: Methods and Results*, volume 4. D. Appleton and Company, New York.
- Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Clarendon Press, Oxford.
- Kim, J. (1989). Mechanism, purpose, and explanatory exclusion. *Philosophical Perspectives*, 3:77–108.
- Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton University Press, Princeton.
- Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, 61(4):343–377.
- Lewis, D. (2000). Causation as influence. *The Journal of Philosophy*, 94(4):182–197.
- Menzies, P. (2009). Mental causation in a physical world. Plenary talk for the Australasian Association of Philosophy 2009 meeting, Melbourne.
- Pereboom, D. (2001). *Living Without Free Will*. Cambridge University Press, Cambridge.
- Pereboom, D. (2002). Robust nonreductive materialism. *The Journal of Philosophy*, 99(10):499–531.
- Quine, W. v. O. (1948). On what there is. *Review of Metaphysics*, 2(5):21–38.
- Sanford, D. (2005). Distinctness and non-identity. *Analysis*, 65(4):269–74.
- Schrenk, M. (2011). Interfering with nomological necessity. *The Philosophical Quarterly*, 61(244):577–597.
- Sellars, W. (1948). Concepts as involving laws and inconceivable without them. *Philosophy of Science*, 15:287–315.

- Shapiro, L. and Sober, E. (2007). Epiphenomenalism — the do's and the don'ts. In Wolters, G. and Machamer, P., editors, *Thinking about Causes: From Greek Philosophy to Modern Physics*. University of Pittsburgh Press, Pittsburgh.
- Shoemaker, S. (1984). Causality and properties. In *Identity, Cause and Mind: Philosophical Essays*, pages 206–233. Cambridge University Press, Cambridge.
- Shoemaker, S. (1998). Realization and mental causation. *Proceedings of the Twentieth World Congress of Philosophy*, 9:23–33.
- Shoemaker, S. (2007). *Physical Realization*. Oxford University Press, Oxford.
- Watkins, M. (2002). *Rediscovering Colors: A Study in Pollyanna Realism*, volume 88 of *Philosophical Studies*. Kluwer Academic Publishers, Dordrecht.
- Wegner, D. M. (2002). *The Illusion of Conscious Will*. The MIT Press, Cambridge, MA.
- Yablo, S. (1992). Mental causation. *The Philosophical Review*, 101(2):245–280.
- Yablo, S. (1998). Does ontology rest on a mistake? *Proceedings of the Aristotelian Society - Supplementary*, 72(1):229–261.
- Yablo, S. (2009). Must existence-questions have answers? In Chalmers, D. J., Manley, D., and Wasserman, R., editors, *Metametaphysics: New Essays on the Foundations of Ontology*, pages 507–525. Clarendon Press, Oxford.