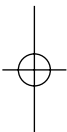
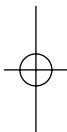




6


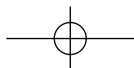
Emergent Causation and Property Causation

Paul Noordhof



When we say that one thing caused another in virtue of certain of its properties, we attribute a case of property causation. A necessary condition for property causation is that an *instance* of the property cited is a cause. This is not a sufficient condition because property causation, as opposed to property instance causation, involves a certain kind of generality. A sufficient condition for property causation is that there should be a law that things with F (a putative property cause) cause things with G; in brief, that the Fs cause the Gs. To fix ideas—the ideas that introduce the problem upon which this chapter will focus—suppose that a subclass of the class of properties identified by current physics, or a future physics sufficiently resembling our own, contains just properties which are property causes in this way. Call these the *narrowly physical property causes*. They are narrowly physical because, presumably, our intuitive notion of a physical property is not exhausted by the properties identified by physics, and certainly not those which are property causes. Suppose, further, that there are properties that stand in a certain highly specific relation of supervenience to the narrowly physical properties, the exact character of which will be set out and defended in section 1. Call these the *broadly physical properties*. A very familiar thought is that if we get the characterization of the supervenience relation right, then the properties classified as either narrowly or broadly physical will capture our intuitive notion of a physical property. Properties that stand in this relationship of supervenience are not emergent properties *relative to the narrowly physical properties*. That does not mean that all emergent properties fail to stand in the relationship of supervenience. There may be some members of the class of narrowly physical property causes that are emergent with regard to other members of the class (for more discussion see, e.g., Hüttemann and Papineau 2005; Papineau forthcoming). Such properties would, of course, trivially supervene upon themselves.

The issue of property causation with which I am concerned is whether there may be any other property causes than the narrowly physical property causes, and the implications of my answer for the proper characterization of emergent causation (if such there be).



Positive answers to the question can take one of two forms. Either it can be argued that there are other property causes because these other properties stand in a certain relationship to the narrowly physical property causes, or an independent analysis can be given of the nature of property causation which serves to classify both the narrowly physical and other properties as property causes. The options, at this point, in part depend on the account of causation, in general, that is favoured. I have argued elsewhere that a counterfactual theory of causation ought to be adopted for particulars, amongst which I include property instances. My aim is to defend its application here and to derive from it an account of property causation.

Thus I shall argue that a univocal independent analysis of property causation can be provided. Appeal to the relationship between other broadly physical properties and those which are narrowly physical provides a justification of the verdicts arrived at by the independent account. It is no part of the correct theory of property causation. Nevertheless, one dimension of the relationship needs to be recognized in order to identify the right kind of generality indicative of property causation. Such an appeal is not required in the characterization of emergent property causation.

In the first section of this chapter, I will provide a characterization of the different relationship of broadly physical properties and emergent properties to narrowly physical properties, and defend it against objections, in particular, that the characterization I favour will not work if a causal theory of properties is true. I will also criticize the alternatives which have been offered as either not allowing for the existence of broadly physical properties or lacking appropriate independence from the key idea to which I appeal. In the second section, I will explain why, given Kim's exclusion argument, the verdicts that a counterfactual theory of property instance causation supplies for broadly physical properties require independent justification where those for emergent properties do not. In the third section, I outline how such a justification may be provided and develop an account of property causation which builds upon it. In the fourth section, I examine the implications of my analysis for the question of emergent causation, focusing in particular upon the question: does all emergent causation involve emergent property causation? My negative conclusion here will enable me to identify a second kind of emergence: emergent non-reductive physicalism. In the concluding fifth section, I briefly consider the consequences of my discussion for two candidates for emergence: phenomenal consciousness and free will.

1. THE PROPER CHARACTERIZATION OF EMERGENCE

Both non-reductive physicalism and emergent dualism involve the idea of property determination. In some sense, both suppose that the instantiation of narrowly physical properties determines the instantiation of the other target

properties: the non-reductive or emergent ones. They differ over whether the instantiation of these other properties involves something genuinely new. Non-reductive physicalists deny this, whereas emergent dualists assert it. The problem is to make sense of when there is something new introduced.

According to the view I favour, a preliminary characterization can be provided in terms of different strengths of strong supervenience. As is familiar, Jaegwon Kim formulated it as follows.

A strongly supervenes on B just in case, necessarily, for each x and each property F in A, if x has F, then there is a property G in B such that x has G, and necessarily if any y has G, it has F. i.e. $\Box (x)(F)(Fx \ \& \ F \ e \ A \ \Box (G) (G \ e \ B \ \& \ Gx \ \& \ \Box (y)(Gy \ \Box \ Fy))$. (Kim 1993 [1984]: 65)

(where A and B are families of properties, the supervening and supervenience-base (or subvenient) properties respectively, 'e' is 'is a member of', and ' \Box ' is the necessity operator with a force to be specified).

Let the family of B properties be the narrowly physical property causes or spatiotemporal arrangements of such properties (I will supplement this shortly). Let the family of A properties be our target properties: mental properties, biological properties, or whatever. Whether one is a non-reductive physicalist or emergentist, it is plausible that the first modal operator should be understood in terms of nomic necessity. Neither position is to be distinguished by the fact that they require underlying supervenience-base properties in all possible worlds. The difference between non-reductive physicalism and emergent dualism lies in the interpretation of the second operator. Non-reductive physicalists should take it to be metaphysical necessity, emergent dualists should take it to be nomological necessity. The intuitive thought is that the reason why emergent dualists reject appeal to metaphysical necessity is that they suppose that some of the target properties determined by narrowly physical property causes are wholly distinct from them, whereas non-reductive physicalists are committed to thinking that they are not.

Of course, both non-reductive physicalists and emergent dualists deny property *identity*. Denying property identity (rather than property instance identity) is what makes non-reductive physicalists non-reductive. However, they add to this thought that even though the properties are not identical, instances of their target properties are, in some sense, nothing more than instances of narrowly physical properties. It is this that the appeal to metaphysical necessity is meant to articulate. The reason why emergent dualists should, at least, appeal to nomological necessity is that they claim that their target properties emerge rather than float free of the narrowly physical property causes. The fact that emergentist dualists need to appeal to fundamental laws to explicate this emergence provides a relatively ontologically robust sense in which the emergence is inexplicable. In brief, non-reductive physicalists believe in M-strong supervenience whereas emergent dualists believe in N-strong supervenience for their target properties

(where M and N indicate the interpretation of the second modal operator: metaphysical or nomological necessity respectively). Others who take a similar line demarcating this position include James Van Cleve (1990: 222) and David Chalmers (2006).

Although the idea has been articulated with respect to a particular characterization of strong supervenience, the latter is not essential to it. A currently (and rightly) popular formulation of physicalism is:

Any world which is a minimal physical duplicate of our world is a duplicate simpliciter of our world. (Jackson 1998: 12)

A minimal physical duplicate is a duplicate in terms of instantiations of narrowly physical properties and stopping right there. Such a world would be a duplicate in terms of instantiations of broadly physical properties as well. However, if emergent dualism were true, then the characterization of physicalism would be false. Instead, the emergent dualist holds that

Any world which is a minimal physical duplicate of our world and has the same laws (where these might include fundamental physico-psychological laws for instance), is a duplicate simpliciter of our world.

The formulation of physicalism characterizes a relationship that should hold in all possible worlds—as we would expect from the appeal to metaphysical necessity in the strong supervenience claim—whereas the second limits it to worlds with the same laws.¹

The favoured characterization of non-reductive physicalism does not rule out the possibility of the supervenience-base properties being relational so as to include features of the environment in which individuals are located, and their interactions with those features. It also does not rule out the possibility that some interactions between instances of mental properties and instances of narrowly physical property causes are on the same level, for instance the macro-level. It should not be assumed that all narrowly physical property causes are micro-properties. Nor does the characterization rule out the possibility that there is an interplay between a number of levels. The important point is that any property that is a potential threat to the truth of physicalism will not be identified by a physics resembling our own. For these properties, if there is a lower level supervenience-base for them, the particular characterization of supervenience I have given explains under what circumstances these properties will be classified as broadly physical as opposed to non-physical emergent properties. As a result, the kind of ontological emergence favoured by Michael Silberstein, who emphasizes the features I have just listed, is best classified as another version of non-reductive physicalism (Silberstein 2006). Silberstein rejects this classification on

¹ For the present purposes, weaker notions of emergence can be classified as types of non-reductive physicalism (for discussion see Chalmers 2006: 252–3; Bedau 1997: 377–9, 393–5; McGinn 1989).

the grounds that the kind of emergence he envisages involves emergent causation. I argue later that emergent causation is compatible with the properties it involves being broadly physical (and hence weakly emergent at best).

Apparent counter-examples to my characterization of emergence in terms of N-supervenience without M-supervenience are causal role properties or second order properties, those possessed in virtue of the fact that a certain causal role is occupied (hereafter, occupant attributing properties). On the assumption that the laws of nature are independent of the properties which are instantiated (though not necessarily the pattern of instantiations), the connection between narrowly physical property causes and causal role properties (say) must be a matter of nomological necessity alone (O'Connor 1994: 96). In different worlds, the very same narrowly physical property causes may have different causal roles in virtue of different laws that hold. Nevertheless, intuitively, causal role properties are not thought of as emergent properties. Functionalism, to take an obvious example, is one way in which non-reductive physicalism may be true (indeed, it is the most commonly cited example). Yet, whether functionalism takes mental properties to be causal role properties or occupant attributing properties, the connection between these properties and narrowly physical property causes is only one of nomological necessity.

It is at this point that I need to qualify my characterization of the supervenience-base of broadly physical properties. The supervenience-base should not just include narrowly physical property causes but also any laws concerning them alone. Once we include these, it will be metaphysically necessary that, given these laws and the narrowly physical property causes instantiated, their causal role will be instantiated. Of course, if we allowed brute physico-psychological laws to be part of the supervenience-base, emergent psychological laws would be necessitated too. The key difference is that there is no reason to suppose that brute physico-psychological laws are part of an intuitive characterization of the narrowly physical. Indeed, if there are emergent non-physical psychological properties, then clearly these laws are not part of the proper characterization of the narrowly physical. Hence, we have a sense in which physico-psychological laws would be explicable by narrowly physical properties and laws about them, if psychological properties were causal role properties or occupant attributing properties, that we would not have if they were emergent (cf. Sober 1999: 142–4). Other potential counter-examples—for instance, if mental properties are environment-dependent—can be accommodated by allowing their supervenience-base to include narrowly physical relational properties.

Another concern arises if a causal theory of properties is true so that the natures of all properties, or at least scientifically fundamental properties, are given by their causal role (e.g. Shoemaker 1980). Then appeal to the contrast between metaphysical and nomological necessity does not appear to be available. According to such a theory, it seems it is not possible for something to be the very same property and yet lack some aspect of its causal role. Suppose that the

coinstantiation of a certain group of narrowly physical property causes yields an intuitively non-physical emergent property. Then, the thought runs, it won't be metaphysically possible for the group in question to be present with the emergent property absent. Yet, the property is supposed to be an emergent one and not a broadly physical property (O'Connor 1994: 97; Wilson 2005: 436–47).

The first point to make is that, if the nature of a property is given by its causal role, it doesn't follow that its causal properties are essential to it or, if you pack essence into nature then you should not assume that properties characterized by causal role alone have that causal role as their nature. Thus a property *P* may have causal role R_1 in world w_1 , R_2 in w_2 , R_3 in w_3 etc. There are a number of different ways in which this point can be developed in face of the obvious objection that these different world-relative roles imply that the property instantiated is different. Suppose, first, that properties are not world bound. They can be instantiated in different possible worlds. Then it is very plausible that, to be instantiated, they must have at least some of their causal role. Nevertheless, it is not clear why they should have all of it. If the laws of our world are holistic, as Jessica Wilson remarks, this is, at best, evidence that the properties of our world are not emergent and, hence, can provide no counter-examples to my way of demarcating emergence (Wilson 2005: 446).²

If the very same property is instantiated in different worlds, it might be wondered what could explain why its causal role is different. We could hardly appeal to different laws if the properties themselves are the basis of laws (see, e.g., O'Connor 2000a: 117–18). The obvious answer is that it, the property, is a bit different. These differences are not sufficient, though, to lead us to conclude that we are talking about a different property. We would say that if we considered two such properties, side by side, in our world (as it were). However, what we would say in our world does not imply that we should say the same when considering a property in our world and one in another possible world. Perhaps the background thought is that, if a property has no other nature than its causal role, then we could not point to something common to the property in each world, as we could if there were a common nature independent of causal role. This betrays a certain presumption as to what must count as common. A certain degree of similarity of position in the causal nexus might suffice. This would be common to each occurrence of the property.

Suppose now we consider properties to be world bound (Heller 1998). The question of whether a particular causal role property may be instantiated in other worlds becomes the question of whether counterparts of that property in other worlds may have a different causal role. There is no reason to reject this possibility. Those who claim that properties are to be understood entirely in terms of their causal roles don't deny that there are worlds with different laws. We may argue that, for any world-bound property, the laws hold in virtue of its

² Unfortunately, for reasons of space, I cannot discuss Wilson's challenge further.

nature. Nevertheless, in different worlds, the counterpart of that property may have a different nature and hence different laws hold. Laws may be intrinsic to properties without being essential to their instantiation in the same way that particulars may have accidentally intrinsic properties (for details on the latter, see Lewis 1986b: 198–209).

A second line of response, to the worry that the truth of a causal theory of properties would undermine the attempted demarcation of emergent property dualism from non-reductive physicalism, is that, if a causal theory of properties is true, there is no reason to be an emergent *property* dualist. Emergent property dualism has two features. First, it holds that properties emerge from narrowly physical property causes. Second, it claims that these emergent properties are not even broadly physical. Given a causal theory of properties, these two features are in tension.

If the nature of a property is determined by all the entities mentioned in the causal role, then the fact that putatively narrowly physical properties have non-physical consequences—instances of emergent properties—impugns their physical status. In which case, it would not be true that narrowly *physical* property causes stand in a metaphysically necessary relationship to emergent properties. If only a subsection of the causal role determines the nature of the narrowly physical properties—presumably their narrowly physical role—then this problem can be avoided. However, it seems to be avoided only at the expense of throwing into question whether the non-physical mental role should be taken to be essential to narrowly physical properties. If it should not, then the first line of response comes into play again.

In addition, it is questionable whether, if a causal theory of properties is true, the grounds on which we might claim that a property is non-physical remain. It is not as if, in these circumstances, they have, or are, phenomenal properties in a damaging non-physical sense. This is not to deny that there are phenomenal properties in the circumstances envisaged. If a functionalist theory of phenomenal properties is correct, then the causal theory of properties will allow that some properties are phenomenal properties without them being non-physical. The point is simply that the motivation for emergent dualism—drawing on a different view about the character of phenomenal properties—is undermined. In which case, there is no need to accept that emergent property dualism is compatible with metaphysically necessary relations between narrowly physical property causes and their emergent properties.

Two cautions are needed. The first is that the argument I have just run does not establish that mental properties are not emergent. The claim is rather that, if they are emergent, this is not because they are distinctively non-physical. As I have already noted, certain narrowly physical properties may be emergent with regard to other narrowly physical properties of that class. My claim is just that there would be no reason to suppose that mental emergent properties are any

different. It is emergent property *dualism* which is threatened. In section 4, I will discuss the nature of emergent physicalism.

The second caution is that the argument is not intended to establish that, if a causal theory of properties is true, then physicalism is true (see, e.g., Shoemaker 1981: 274–8). For all I know, non-physical worlds of causally characterized properties are possible. Rather, my more limited aim is to establish that, if we have a world of pervasively instantiated narrowly physical property causes and the causal theory of properties is true, it is unclear how emergent property dualism is true.

Philosophers, convinced that the appeal to metaphysical necessity in the characterization of supervenience is mistaken, stick with nomological necessity and attempt to supplement it with the requirement that, in the case of non-reductive physicalism, we have an explanatory relationship between narrowly physical property causes and broadly physical properties whereas, in the case of emergent properties, the relationship is not explanatory. Matters turn on the favoured type of explanation. The two standard approaches take broadly physical properties to be functionalizable or macro-properties respectively. According to the former, favoured by Jaegwon Kim, broadly physical properties are those which are instantiated if another property occupies a certain causal role (see, e.g., Kim 1998: 100–1). According to the latter, defended by Timothy O'Connor and co-writer, Hong Yu Wong, broadly physical properties are structural properties in which the proper parts of particulars which possess them have properties, not identical with the structural property, jointly standing in relation R.

The problem with the first option, appealing to second order causal role properties, is that it is unclear whether this supplies us with an account of non-reductive physicalism in which mental properties exist. It seems more plausible to suppose that we have mental concepts, the concept of some property or other that occupies a certain causal role, which we apply to particulars partly in virtue of the property that occupies the role. Indeed, this is the conclusion that Kim draws in the development of his position (Kim 1998: 104; see also Wilson 2005: 453).

The problem with the second option, appealing to the idea of structural properties, concerns variable realization. O'Connor and Wong claim that if a particular possesses the relevant parts with properties standing in the appropriate relation R, then they possess the structural property. There is *nothing more* to the possession of the structural property than that (O'Connor 1994: 93; O'Connor and Wong 2005: 663). The question is how to understand this 'nothing more'. In certain cases, identity seems plausible. For instance, possession of the structural property of being methane is nothing more than instances of the properties of being carbon and hydrogen atoms in a certain arrangement R.

In the case of the property of being an earthquake, a similar approach does not seem to work. There is no particular arrangement of parts with properties which holds for every case of earthquakes. There are a vast variety of different

ways in which there could be an earthquake. Moreover, the more we seek to characterize what might be common to all cases of earthquakes, the more the properties with which we characterize this commonality are unlikely to be narrowly physical property causes. The proposal thus faces a dilemma. If you go highly specific and seek to characterize everything in terms of narrowly physical property causes, then we need some other way of characterizing why the property of being an earthquake, as opposed to instances of the property of being an earthquake, is a broadly physical property. On the other hand, if you go highly general, you might be able to capture what is common to all cases of the property of being an earthquake (let us suppose), but now you open up a gap between the narrowly physical property causes and the more general specification.

It might be argued that the property of being an earthquake is a broadly physical property because all of its instances involve narrowly physical properties. However, this will not do as an answer by itself. If you claim that two properties have the same instance, then nothing has been established as to whether the properties in question are physical. Instead, the instance may involve both physical and non-physical properties. If it is claimed that instances don't combine physical and non-physical properties, we need to know why. Whatever answer is given at this stage is the answer to our question of what is involved in 'nothing more' and not the claim that the putative broadly physical property only has instances that are physical. For example, if it is said that two properties do not have the same instance unless one of them metaphysically necessitates the other, then that is the answer to our question. Unconstrained talk of identity of instances cannot yield the right results. Suppose, for example, that narrowly physical property causes nomically necessitate certain emergent properties. We could not establish that the emergent properties were physical rather than non-physical simply by saying that the same instance has these nomically related elements. Our defence of the claim that they have the same instance requires discussion of the relationship between the properties themselves and whether this relationship is compatible with them sharing instances. Certainly, if it is not metaphysically possible for there to be an instance of the property of being an earthquake without it being an instance of an arrangement of narrowly physical property causes, then this shows something about the nature of the property of being an earthquake. But now we simply have an assertion of the relation of supervenience which was being eschewed. For these reasons, O'Connor and Wong's approach doesn't seem to avoid the problem to which the appeal to a metaphysically necessary supervenience relation was the solution.

Once appeal to metaphysical necessity is in play, it is reasonable to consider whether we need anything else. We would have captured the idea that emergent properties are, in some sense, inexplicable by saying that their instantiation requires a fundamental physico-psychological law. The idea that there is a metaphysically necessary connection between arrangements of narrowly

physical properties and broadly physical properties is a placemaker for various kinds of explication that might be provided of the connection. There seems no reason to limit ourselves to certain types or even claim that we will be able to grasp the kind of explication that is at the basis of the metaphysical necessity—for instance, we might be cognitively closed to it (McGinn 1989).

I conclude that the distinction between broadly physical properties and a strong kind of emergence is best made out by a difference in the strength of the second necessity operator in the formulation of supervenience and that no other account promises to do better without appealing to the same materials.

2. A COUNTERFACTUAL ANALYSIS OF PROPERTY INSTANCE CAUSATION

The way in which I have distinguished broadly physical from non-physical emergent properties has consequences for the development of a counterfactual analysis of property instance causation and, thereby, of property causation. For the purposes of the present discussion, I am going to make two simplifying assumptions. The first is that, subject to a successful response to the objection below,

p_1 is a cause of p_2 if (1) if p_1 were not instantiated, then p_2 would not be instantiated and (2) if p_1 were instantiated, p_2 would be instantiated (where ' p_1 ' and ' p_2 ' are property instances and they are distinct existences in a sense to be made clearer below).

So I am setting aside issues of indeterminism for which appeal to chance raising counterfactuals would be necessary and, by only providing a sufficient condition, not engaging with the question of redundant causation (see, e.g., Noordhof 1999c).

The second simplifying assumption is that, for our purposes, Lewis's analysis of, and similarity weighting for, counterfactuals is correct. Thus

A counterfactual 'If it were that A, then it would be that C' is (non-vacuously) true if and only if some (accessible) world where both A and C are true is more similar to our actual world, overall, than is any world where A is true but C is false. (Lewis 1986a [1979]: 41)

with similarity of worlds measured by the following conditions:

- (A) It is of the first importance to avoid big, widespread, diverse violations of law.
- (B) It is of the second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.
- (C) It is of the third importance to avoid even small, localized, simple violations of law.
- (D) It is of little or no importance to secure approximate similarity of particular fact, even in matters which concern us greatly. (Lewis 1986a [1979]: 47–8)

In fact, the similarity weighting needs some reform but the ways in which it needs reform should not touch the discussion to follow.

With these assumptions in place, we can see the problem for a counterfactual analysis. Let ep_1 be an instance of an emergent property and $A(p_1, p_2, p_3, \dots)$ be an arrangement of narrowly physical property causes from which ep_1 emerges. It seems possible that ep_1 is not a cause of e but $A(p_1, p_2, p_3, \dots)$ is. The case seems to be one in which a cause, $A(p_1, p_2, p_3, \dots)$, has two effects ep_1 and e without it following that one of the effects is a cause of the other. The standard treatment of this in a counterfactual analysis turns on whether a backtracking conditional between ep_1 and $A(p_1, p_2, p_3, \dots)$ holds. There seems no more reason to allow for the possibility of backtracking counterfactuals being true in the particular case of emergent properties and their N-supervenience-bases than in standard cases of two effects. We obtain more perfect match by retaining $A(p_1, p_2, p_3, \dots)$ and violating the law to ep_1 than by removing $A(p_1, p_2, p_3, \dots)$, keeping the law, and securing additional perfect match by covering up the traces of $A(p_1, p_2, p_3, \dots)$ in the future. So, by the similarity weighting, the closest worlds will not be worlds in which, when ep_1 is absent, $A(p_1, p_2, p_3, \dots)$ is absent. There seems to be no possibility of the counterfactual analysis of property instance causation falsely concluding that there is causation when dealing with emergent properties.

The situation is different for broadly physical properties. Let bp_1 be an instance of a broadly physical property, and $A(p_1, p_2, p_3, \dots)$ the arrangement of narrowly physical property causes which constitutes its M-supervenience-base. At its most basic, this may just be the instantiation of a single narrowly physical property—the formulation is just for generality. There are two types of broadly physical properties we need to consider: occupant attributing properties and other broadly physical properties which are putative occupants of a causal role. For each, the challenge breaks down into two components. First, can there be cases in which $A(p_1, p_2, p_3, \dots)$ is efficacious and bp_1 not? Second, if so, can a counterfactual theory capture this? I will consider these questions in reverse order since it will usefully narrow the discussion.

Before I do this, though, I will make a third simplifying assumption which has the effect of ruling out one possible line of escape for the counterfactual theorist. When we consider what would be the case were bp_1 to be absent, one way of retaining the presence of e is to suppose that there might be some replacement property(ies) instantiated, bp_1^* perhaps or simply some $A(p_1, p_2, p_3, \dots)^*$, which still causes e . We don't have a case of redundant causation exactly because these replacement properties are not actually instantiated. Nevertheless, they would be in the envisaged counterfactual circumstances. Call this *close-world redundant causation*. There are ways to deal with this within the counterfactual framework (see, e.g., Ganeri, Noordhof and Ramachandran 1996, 1998; Ramachandran 1997; Noordhof 1999c). The important point is that they

should offer no succour to the counterfactual theorist with regard to the present problem.

With the assumption in place, we should conclude that if the occupant attributing properties were not instantiated, then e would not occur either. That would be so, even if $A(p_1, p_2, p_3, \dots)$ were still present. The changes in the laws required for the absence of the occupant attributing property would also make $A(p_1, p_2, p_3, \dots)$ not a cause of e . So the challenge to the counterfactual analysis of property instance causation is direct. If occupant attributing properties are inefficacious with regard to e , the counterfactual analysis will not capture this verdict. For other broadly physical properties which don't have, as part of their supervenience-base, laws relating narrowly physical properties—macro-properties might be an example—their absence would mean that $A(p_1, p_2, p_3, \dots)$ is absent. According to my account of broadly physical properties, there is a metaphysically necessary relationship between instances of $A(p_1, p_2, p_3, \dots)$ and these other broadly physical properties. So, once more, e would not occur. This time because its cause, $A(p_1, p_2, p_3, \dots)$, is absent.

It seems that the counterfactual theory of property instance causation cannot allow that broadly physical properties of either sort fail to have efficacy in the circumstances envisaged. Yet these verdicts are, to put it mildly, contested. Here is how Jaegwon Kim expresses the worry (e.g. Kim 1998). I adapt it for the particular account of the relationship between broadly physical properties and narrowly physical property causes I defended in the previous section. First, it is assumed that physics is complete, in our terminology the sufficient causes of all instances of narrowly physical property causes are other narrowly physical property causes. Suppose that $A_1(p_1, p_2, p_3, \dots)$ is causally sufficient for $A_2(p_{100}, p_{101}, p_{102}, \dots)$ and that the first is the supervenience-base for bp_1 and the second is the supervenience-base for bp_2 . Assume, for the sake of argument, that bp_1 is a cause of bp_2 ; bp_1 causes bp_2 either directly or by causing $A_2(p_{100}, p_{101}, p_{102}, \dots)$. If it causes bp_2 directly, then either $A_1(p_1, p_2, p_3, \dots)$ is insufficient for bp_2 by causing $A_2(p_{100}, p_{101}, p_{102}, \dots)$ or bp_1 is an overdetermining cause. If bp_1 causes bp_2 indirectly by causing $A_2(p_{100}, p_{101}, p_{102}, \dots)$, then the same choice holds regarding $A_2(p_{100}, p_{101}, p_{102}, \dots)$. It is implausible to suppose that there should be systematic—since the reasoning is entirely general and the situation envisaged widespread—overdetermination in either of these ways. Therefore, bp_1 is inefficacious.

At this point, the rejection of all kinds of interlevel causation can seem a tempting defence for the counterfactualist. A key idea is that causation must involve the right level of generalization and not simply determination (e.g. see Gibbons 2006). A sophisticated development of this strategy holds that different levels are associated with different patterns of close-world redundancy (Menzies this volume). Thus, if we consider what would be the case if a certain instance of a mental property were absent, we must suppose that all of its M -supervenience-bases are absent. Whereas, if we consider what would be the case if a certain

M-supervenience-base is absent, then (arguably) another M-supervenience-base of the property would be there instead. So we might be led to conclude that the mental property is efficacious and its M-supervenience-base is not.

It seems to me both that there are difficulties in implementing this strategy and problems with it as a dialectical move. On the implementation side, the kind of properties that distinguish a level often have their character specified in terms of a causal role. For instance, actions are partly characterized in terms of their causal history, part of which will include an intention. The standard response to the claim that entities so characterized cannot stand in causal relations, because they fail to be distinct from each other, is that they can because of the existence of lower level entities not specified in that fashion. Thus, intention and action are related as cause to effect because of their narrowly physical properties (e.g. Davidson 1963: 14–17). If the efficacy of the lower level is not allowed to transmit to the higher level, we have no explanation of why this lower-level causation makes it true to say that intentions cause actions. If it is allowed to transmit, then it is unclear why the lower-level physical properties cannot be counted as causes of actions as well, so giving rise to the familiar concern.

As a dialectical move, the worry about the denial of interlevel causation is that, even if causation is not simply a matter of determination, this is a necessary condition for causation. Suppose we give the term ‘cause’ to the deniers of interlevel causation and, thereby, accept that narrowly physical property instances are not causes of higher-level properties and vice versa. Deniers of interlevel causation must still accept that there is a competition of determination. If it can be shown—and this is what Kim’s argument purports to do—that determination resides at the level of narrowly physical property instances, then it is still not plausible that the higher-level properties count as causes of entities at their own level. If, on the other hand, determination transmits, then there will be no problem. But in those circumstances, we could go straight to GO with a determination account of causality, and property instance causation in particular.

A proponent of workers’ rights might also protest that the honest toil of the narrowly physical properties should not be impugned because, if the work had not been done by one of them, it would have been done by another, and should not accrue to the mental property just because it would be there whichever narrowly physical property was in play. According to such a picture, it would seem that something would only be credited with the work it did if there were nothing else on hand to do it in its place.

The problem that my preferred way of distinguishing between non-reductive physicalism and emergent property dualism presents, then, may be put like this. The counterfactual theory of property instance causation and Kim’s argument are in conflict over the cases. The metaphysically necessary relationship between the supervenience-base of broadly physical properties and these broadly physical properties is a plausible case of counterfactual dependence spuriously indicating

causal dependence. Successful defence of a counterfactual theory of property instance causation requires a treatment of this concern together with a diagnosis of what is wrong with Kim's argument. A similar challenge would be faced by conditional chance-raising accounts of causation (see, e.g., Sober 1999: 145–8).

3. AN ACCOUNT OF PROPERTY CAUSATION

As I have already noted, an account of property causation breaks down into two components: an account of property instance causation and the element of generality which grounds the claim that it is the property and not just a particular instance of it that is efficacious. If a counterfactual analysis of property instance causation is appropriate for the first component, we need an explanation of how, with some qualifications, if the M-supervenience-base properties of a broadly physical property are efficacious, then the broadly physical property is efficacious. As we saw in the previous section, the counterfactual analysis will pronounce most of them to be so. Without such an explanation, its credibility is undermined.

To this end, I wish to defend a qualified version of the principle that

(TC) If an instance of $A(p_i, p_{i+1}, p_{i+2}, \dots)$ is a cause of e and $\Box_m(x)(A(p_i, p_{i+1}, p_{i+2}, \dots) \supset x \supset bp)$, then the instantiation of bp is a cause of e .

I have labelled it (TC) for transmission of causality principle since it is a response to the intuition that there is no doubting the efficacy of narrowly physical property causes. The question is whether there is any further property causation. In fact, I question whether this is the appropriate way to look at things. The more neutral way of posing the issue is whether there can be harmony regarding the attribution of efficacy at different levels. Since this does not alter the character of the argument to follow, I have suppressed such qualms.

I have a two part defence of the principle. The first is a generalization from intuitively efficacious broadly physical properties. The second is an accusation of asymmetry.

If we begin with mental properties and are convinced that they are broadly physical because we are physicalists, then the combination of Kim's argument plus residual worries about how mental properties fit with the narrowly physical world can convince us that they are inefficacious. They just don't seem the right kind of thing to have an impact upon the instantiation of narrowly physical properties. If we begin instead with earthquakes, rivers or glaciers, match lightings and other events characterized by macro-properties, Kim's argument seems much less compelling. Instantiations of all of these properties are metaphysically necessitated by some arrangement of narrowly physical property causes and yet they are efficacious. This provides some support for the transmission of causality principle.

Simple cases of determinable and determinate properties provide us with additional support. If my eleven-stone weight caused the chair to break, it seems churlish to deny that my having some weight or other did so, too. Obviously, in this case, the former seems more adequate than the latter (a matter I will come back to in a moment). There are examples in which the more determinable property captures what is required where a determinate seems superfluous. According to the myth, bulls are enraged by red capes. Presumably, then, the fact that this particular cape is scarlet is beside the point. If there is no univocal proclamation in favour of the determinate, then it is plausible that determinable properties may still be efficacious in circumstances in which they are inadequate but necessary. If there is no univocal proclamation in favour of the determinable, then it is plausible that determinate properties may be efficacious while still having superfluous features.

One response to this line of reasoning is to reject it on the grounds that focus on mental properties has simply helped us to question an unjustified orthodoxy with regard to the other broadly physical properties mentioned, being an earthquake, a river, or a match lighting. This is not a response that Kim emphasizes. Instead, he seeks to introduce a block to the generality of the argument, so isolating mental properties from their unproblematic broadly physical potential cousins. There are some cases in which supervening properties are inefficacious because efficacy resides in their supervenience-base and there are some which are efficacious: the micro-based properties. These he characterizes as follows.

P is a micro-based property just in case P is the property of being completely decomposable into nonoverlapping proper parts, a_1, a_2, \dots, a_n such that $P_1a_1, P_2a_2, \dots, P_n a_n$ and $R(a_1, a_2, \dots, a_n)$. (Kim 1998: 84)

Intuitively efficacious broadly physical properties turn out to be micro-based, mental properties sadly do not.

Kim's response has caused puzzlement. The argument at the end of the previous section seemed to rely upon the following facts. First, the supervenience-base properties necessitated broadly physical properties. Second, instantiations of supervenience-base properties were sufficient causes of other supervenience-base properties. These facts hold just as much for the relationship between arrangements of narrowly physical property causes and micro-based properties as they do between the former and other supervening properties (Noordhof 1999b: 109–14). I take it that an object O's possession of a micro-based property is distinct from there being entities a_1, a_2, \dots, a_n such that $P_1a_1, P_2a_2, \dots, P_n a_n$ and $R(a_1, a_2, \dots, a_n)$ for two reasons. First, the micro-based property is a property of the object rather than a series of properties of the entities a_1, a_2, \dots, a_n , together with a relation between them. Second, micro-based properties may be variably realized: the property of being an earthquake would be a good example. Because of this, I don't know what to make of Kim's response that his argument does not apply to micro-based properties because the idea of a micro-based property

having a micro-base is obscure and there is no relation of determination (as opposed, perhaps, to identity) between there being entities a_1, a_2, \dots, a_n such that $P_1 a_1, P_2 a_2, \dots, P_n a_n$ and $R(a_1, a_2, \dots, a_n)$ and the micro-based property (see Kim 1999a: 116–17).

Without a block, it seems to me that we are more confident that properties such as being an earthquake are efficacious than we are about the principles behind Kim's argument. One way of supporting this position is to argue that we need to refine our notion of the causal completeness of physics to allow that it is not impugned by certain other non-narrowly physical property instances being counted as causes. For instance, we might say that physics is complete partly because its proprietary properties necessitate broadly physical properties. The efficacy of bp accrues to $A(p_i, p_{i+1}, p_{i+2}, \dots)$ because of this fact. Even if physics provides complete coverage and focuses on the very small (both contentious claims), there is no reason to insist that this should mean that efficacy is always located first and foremost at the narrowly physical level and then must be spread outwards as it were. Alternatively, we may argue that instances of broadly physical properties just identify a component of the causal relationship between narrowly physical property instances or have as parts instances of narrowly physical causally related properties. Regarding the first of these two possibilities, there is no reason to suppose that every such component should be identified by physics with its own distinctive explanatory aims. All of these options involve explanatory burdens that are not easy to meet but not of the character that threatens the plausibility of this line of response. We shouldn't abandon a philosophical strategy, anymore than we should abandon the search for a cure of a disease, because it is difficult.

Another related way of supporting the recommended position focuses on refining our understanding of overdetermination. Overdetermination is not simply a matter of there being two distinct causes of certain target effect. Two property instances overdetermine an effect only if they are independent causes where this is understood to mean that either one would cause the effect without the other. Two property instances would be genuine causal competitors only if they are independent causes in this sense (Bennett 2003: 476–92; Noordhof 1999a: 305). My characterization of broadly physical properties in terms of M -supervenience rules them out as independent causes of the effects caused by arrangements of narrowly physical properties and hence denies that there is genuine causal competition.

It is not open to opponents simply to resist this refinement of overdetermination and stick with the original one. Everybody agrees that two links in the same causal chain are not competitors and don't overdetermine a certain effect. So we need some characterization of why there is no overdetermination in this case. A flat-footed response is to insist that they are not overdetermining causes because they bring about the effect by the same causal chain. However, this is not

right. The earlier property instance in a causal chain involves connecting causal elements that the later does not.

The next move is to say that the two causal chains should not have wholly overlapping parts (presumably at the section of the chain leading up to the effect). However, two events may cause an effect by two chains with a wholly overlapping part leading up to the effect if either event would not have been able to cause the effect without the other. Imagine that the effect requires vigorous stimulation which can only be brought about by the events in concert. The idea we need is that one chain leading from a cause to the effect could have brought about the effect in the absence of the other, that is, that they are independent causes. However, this is precisely what we don't have in the case of narrowly physical property instances and the broadly physical properties which supervene upon them. Thus a necessary condition for genuine overdetermination is not met.

The second part of my strategy regarding the defence of the transmission of causality principle was the charge of unmotivated asymmetry. The charge concerns the apparently *initially* different attitudes taken to, for instance, macro-causal relations and other macro-properties. To begin with, the existence of instances of macro-properties is not supposed to be in doubt. They are taken to supervene upon arrangements of instances of narrowly physical micro-properties. Afterwards, when their efficacy is impugned, the existence of instances of macro-properties may be denied but it is important to recognize that this is not the initial situation. It is agreed on all sides that instances of micro-causal relations exist. In which case, the instances of macro-causal relations should supervene upon an arrangement of them in the same way. To deny this seems an unmotivated asymmetry. What is so strange about the macro-causal that its instances should not be constituted from instances of micro-causal relations in the same way that instances of macro-properties in general are constituted from instances of narrowly physical micro-properties? The assumption seems to be that *if* the micro-causal relations fix what must happen, then the macro-causal relations have nothing further to contribute. Yet if instances of micro-causal relations do constitute an instance of a macro-causal relation, then the macro-causal relation has already made its contribution and to demand it again is to be unnecessarily insistent.

We have seen that there is one major class of cases in which the transmission of causality principle is defensible. The efficacy of determinable properties—my other illustration of the intuitive force of the transmission principle—has been questioned. The main charge is that determinate properties have all the causal powers attributed to determinable properties and hence there is no reason to take the latter to exist. It appeals to what has been called the 'subset' account of realization. An object that has the *determinate* property of being a sphere of a certain size has all the causal powers that accrue to it as a result of that property. Obviously if the object is a sphere of a certain size, it is also simply a sphere. Yet,

as the reasoning goes, the latter property is not required by the object to confer the causal powers distinctive of being simply spherical. These causal powers are simply a subset of the causal powers of being a spherical object of a certain size. What confers no causal powers does not exist (Gillett and Rives 2005: 490–1).

The reasoning appears questionable in two respects. First, it seems tacitly to rule out the possibility that the instantiation of a determinate property has some of its causal powers *because* it involves the instantiation of a determinable property. This seems to stem from an antecedent commitment to the non-existence of the determinable. In which case, the argument can hardly hope to establish in non-question-begging fashion that determinables do not exist. If instantiations of determinate properties have the determinable's subset of powers by involving their instantiation, then there will be no double-counting of causal powers (a charge made by Gillett and Rives (2005: 486–7)). It is common ground that there is a subset of causal powers and, by being a subset of causal powers, we have a genuine resemblance between objects which possess that subset. That seems a good basis to conclude that determinable powers exist even if the subset view is correct.

Second, in fact, it is not clear that instances of determinable properties have causal powers which are simply a subset of an instance of the determinate property that realizes them. Suppose that there is a circular hole of 4 cm width. A sphere 4 cm or less can go through it, a cube would have to be around 2.8 cm or under to go through. The property of being a sphere has causal powers, then, that the property of being a 5 cm sphere does not, namely passing through the hole for a range of sizes at or below 4 cm.

The point may be made even more plausibly with regard to particular ways in which the property of being spherical may be realized, which are not appropriately characterized as determinates of the determinable being spherical. For example, suppose a spherical rock travels through a window at speed leaving a spherical hole in the glass (along with cracks spreading outwards). A spherical piece of cotton wool has no such effects. The property of being spherical has causal consequences the property of being a spherical piece of cotton wool does not, namely that the former is a cause of the spherical character of the hole. Such cases only appear problematic if you assume that the powers of a determinable must be derived from a determinate rather than simply be due to the determinables in their own right.

A natural response is to say that, although a determinable property doesn't have a subset of the causal powers of its determinate realizing property on a particular occasion, it can be thought of as having a disjunction of subsets of the causal powers, one subset for each of its determinates (Gillett and Rives 2005: 502, fn. 10). That means that, for a particular instantiation, the determinable property does not have the causal powers it has when instantiated with another of its determinates.

This is fine as a position but it cannot constitute an *argument* against the existence of determinables and, in particular, not a causal argument against their existence. It presumes that there is no reality to a determinable other than its instantiation by a particular determinate of it. Standardly, an object or property has its causal powers regardless of whether it is in circumstances conducive to their manifestation. Applying this thought to the case in hand, the property of being spherical has the power to go through spherical holes for a certain range of sizes or make spherical holes in windows regardless of whether, on a particular occasion, it cannot because it has too great a size or is a characteristic of cotton wool. The claim can only be dismissed if it is thought that the existence of the determinable on a particular occasion is simply the determinate in which it is instantiated. Naturally enough, with this assumption, it will seem obvious that determinables have no distinctive powers free of their determinates and hence that determinables don't exist.

Of course, determinables require some determinate or other to be instantiated. This does not mean that, given all the determinates of a world, we have explained all the causal powers in that world. A world does not involve instantiations of all a determinable's determinates. Even if it did, we would still lack an explanation of how, for a particular instantiation of a determinable, it has causal powers which outstrip those of its determinate. The fact that there are other determinates with different causal powers seems to have no impact upon the causal powers of the particular determinate in question.

If determinables don't exist, I do not pretend to have shown they do. The argument of the last few paragraphs rested upon the assumption that they did exist and considered what followed from that. Nor, then, would we have a counter-example to the principle of the transmission of causality. Efficacy does not transmit to entities that do not exist (of course). My claim is simply that if there is a metaphysically necessary relationship of the required type between instantiations of properties (each of which exist), then the necessitated will be efficacious if the necessitating are.

The support I have produced for the causal transmission principle does not have the upshot that all occupant attributing properties are efficacious. The proper treatment of this type of case divides into two, depending upon whether or not the occupant attributing properties are part of a powers ontology. Suppose, first, that we are not considering properties for which a powers ontology holds. The narrowly physical properties have a causal role to play, settled by the laws that hold. These laws don't M-supervene upon which properties are instantiated (though the pattern of properties that are instantiated may determine which laws hold) (for further discussion, see Noordhof 1997). In those circumstances, the occupant attributing properties are not M-supervenient upon the narrowly physical properties (or particular spatio-temporal arrangements of them). Of course, they would M-supervene upon narrowly physical properties plus the laws (or the particular patterns of properties, which determines that the laws

hold if a regularity view of laws is adhered to). However, we should simply exclude this characterization of the M-supervenience-base. So, one qualification to the transmission of causality principle is that the m-supervenience-base should include neither laws so understood nor the patterns of properties which settle which laws hold.

Suppose now we consider what we should say about occupant attributing properties in a powers ontology. There are at least two options to consider here. If the laws M-supervene upon particular narrowly physical properties (say) then they will, in effect, be part of the supervenience-base for occupant attributing properties. However, this should not threaten the efficacy of such occupant attributing properties. In such circumstances, the transmission of causality principle simply reflects how combinations of causal role at one level—the narrowly physical—relate to the instantiation of causal roles at another level—the mental. So the qualification I made earlier about how the laws should not be part of the M-supervenience-base should be understood to concern only laws which themselves don't M-supervene upon which properties are instantiated. The principle so formulated provides a constraint upon how we should understand the idea that causes and effects must be distinct existences.

In the previous section, I explained how, even if one is committed to a powers ontology, one need not suppose that, in every world, if a property is instantiated, a particular causal role (and hence particular laws) are instantiated. In that case, even according to the powers ontology, laws do not M-supervene upon which properties are instantiated. However, in that case, the occupant attributing properties do not M-supervene upon narrowly physical properties either. Putting these ideas together, we have the following general treatment. Occupant attributing properties are not efficacious if they M-supervene upon a supervenience-base which includes laws which, themselves, fail to M-supervene upon which properties are instantiated.

Another qualification to the transmission of causality principle,

(TC) If an instance of $A(p_i, p_{i+1}, p_{i+2}, \dots)$ is a cause of e and $\Box_m(x)(A(p_i, p_{i+1}, p_{i+2}, \dots)x \rightarrow bp)$, then the instantiation of bp is a cause of e ,

is needed. It arises because the simple relation of metaphysical necessitation does not ensure that the bp has anything to do with the efficacious part of the supervenience-base property. Suppose that $A(p_i, p_{i+1}, p_{i+2}, \dots)$ is a cause of e , that $A(p_k, p_{k+1}, p_{k+2}, \dots)$ (where $k > i$) is not and further that $A(p_k, p_{k+1}, p_{k+2}, \dots)$ metaphysically necessitates bp^* (where $bp^* \neq bp$). Then $A(p_i, p_{i+1}, p_{i+2}, \dots) \& A(p_k, p_{k+1}, p_{k+2}, \dots)$ metaphysically necessitates bp^* . Nevertheless, we would not want to conclude that bp^* was efficacious. Intuitively, the efficacy of the supervenience-base had nothing to do with what was required to instantiate bp^* .

For this reason, we must impose a condition upon $A(p_i, p_{i+1}, p_{i+2}, \dots)$ for it to transmit its efficacy, namely that it should be the minimal supervenience-base

for the instantiation of the property to which the efficacy is being transmitted. As a first stab at this condition, we might hold that B is not the minimal supervenience-base of a property P if P is still instantiated if not B but B⁻ (where B⁻ is B without the instantiation of certain properties and no additional properties are instantiated except those which are metaphysically necessitated by the non-instantiation of these other properties) (Noordhof 1999a: 307). No doubt there may be problems with this proposal as it stands, but the general idea, it seems to me, holds good. Even if an analysis is not possible and we just have to take the idea of a minimal supervenience-base as a primitive, this does not seem to be particularly damaging. For instance, it doesn't seem determined by facts about when the supervening property is efficacious.

With these qualifications in place, we now have a defence of the verdicts that the counterfactual analysis of property instance causation provided for broadly physical properties. However, this does not complete the story. We need to explain how this may be turned into an account, not of property instance causation, but property causation. Elsewhere, and in response to the rejection of interlevel causation, I have explained why it is not productive to consider property causation to be a more refined version of property instance causation in which the putative property causes are somehow more appropriate/adequate for the effects (e.g. Noordhof 1999a: 303–7). Instead, as I have already remarked, an element of generality is needed. This is reflected by the fact that, for the narrowly physical properties, appeal to laws was required. When we turn to properties that supervene upon narrowly physical properties, I think this requirement needs to be relaxed. In its place, I have recommended the following

F is a property cause of G only if each minimal M-supervenience base of F is such that all its instantiations would cause (or in the case of indeterminism, raise the probability of) an instantiation of one of the minimal M-supervenience-bases of G if they were in some causal circumstances C—where C may vary for each kind of supervenience-base. (Noordhof 1999a: 307)

Thus if F has three types of minimal supervenience-bases then each of these should always cause Gs with a certain minimal supervenience-base in a specified set of circumstances. It is plausible to claim that the property is a cause because the identified relationship holds for all the supervenience-bases. Thus there is no reason for supplanting such property causation by property causes which are simply the minimal supervenience-bases in question.

If my characterization of emergent properties is correct, then the condition for property causation in the case of broadly physical properties will not apply to them. Instead, there will be fundamental laws which relate the emergent properties to a particular type of effect. The situation is no different than for narrowly physical property causes. Although the emergent properties supervene in some sense upon narrowly physical property causes, they do not supervene in the way that I require for the condition to apply. Thus the requirement for emergent

property causation is commensurately stronger. It does not allow that there is emergent property causation if the properties upon which emergent properties N-supervene satisfy the corresponding condition formulated in terms of N-supervenience-bases. The intuitive reason for this is that the question of whether a particular emergent property is efficacious is independent of the conditions for its instantiation (although these may be part of the causal conditions in which the emergent property operates).

Some may hold that there is a need for further refinement motivated by the thought that property causation requires not simply generality (the property of being a hammer always causes nails to go in) but also precision (it's not really being a hammer but being an object of a certain weight and resistance that is required for the nails to go in). I have defended the approach above against this line of objection elsewhere so I won't go into it here (see Noordhof 1999a, 2006).

4. EMERGENT CAUSATION

I have distinguished emergent properties from broadly physical properties and explained how a counterfactual analysis of property instance causation can be defended, and developed into an account of property causation. The basic idea is that the favoured analysis should be plugged in where mention of cause occurs in the generality condition supplied at the end of the last section. One way of understanding emergent causation would then be to say that emergent causation is simply causation by emergent properties. However, matters are not so straightforward.

A second, and perhaps central, way to understand emergent causation is as causation which is in some way novel or unpredictable when we consider the causal interactions from which it emerges. Making this more precise is difficult. It should not be thought of as what would strike us as novel because there are possible cases of emergent causation which would fail to do so.

There are two ways in which $A(p_i, p_{i+1}, p_{i+2}, \dots)$ may have consequences which fail to be a result of the laws that govern the instances of narrowly physical properties which make up the arrangement. The first way is if the arrangement has an effect that is neither the result of all the individual causal relations which the instances of properties would have outside of that arrangement, nor the result of the same laws that govern the individual causal relations applying to that arrangement (for further discussion see Noordhof 2003: 90–3). There is no puzzle, for instance, about the baseball's breaking of a window just because the breakage would not occur as a result of all the individual causal relations between elements of the baseball and elements of the glass. The same laws are at work. Here, there seems a relatively straightforward connection between apparent novelty and emergent causation. There is emergent causation because there are *emergent consequences* indicating the operation of different laws.

The second way in which there may be emergent causation is if the arrangement *fails* to have an effect which, taking into account all the individual causal relations and/or the laws concerning them, we would expect it to have. Here we don't so much have a case of emergence as *submergence*. If an arrangement has a submerged consequence, it does not follow that the expected effect fails to occur. An emergent property of the arrangement may cause it to occur. In such circumstances, we might formulate law statements concerning narrowly physical properties oblivious of the fact that arrangements have submerged consequences for which the presence of emergent properties compensates. If $A(p_i, p_{i+1}, p_{i+2}, \dots)$ is sufficient for the emergent property, then there may be no obvious need to mention the presence of the emergent property. The occurrence of the effect subsequent to the instantiation of $A(p_i, p_{i+1}, p_{i+2}, \dots)$ may seem in no way surprising. Nevertheless, we would have a case of emergent causation in the first sense. In these circumstances, the arrangement would not so much fail to have an effect as fail to have it in the way we assume.

The point has application against Kim's argument that such emergent properties will be epiphenomenal. He claims that, if $A(p_i, p_{i+1}, p_{i+2}, \dots)$ is sufficient for an emergent property to be instantiated, it is sufficient for anything which the instance of the emergent property causes. Hence, he concludes, the emergent property has no causal work to do (Kim 1999b: 32–3; see also Sober 1999: 139). The conclusion does not follow. The efficacy of the emergent property is revealed by the fact that, if the emergent property were not present, then the target effect would not occur even though $A(p_i, p_{i+1}, p_{i+2}, \dots)$ is still present. This is quite compatible with allowing that, given that $A(p_i, p_{i+1}, p_{i+2}, \dots)$ is sufficient for the instantiation of the emergent property, it will also be sufficient for the target effect. This is simply a generalization of a familiar point, that the sufficiency of preceding links in the causal chain does not threaten the efficacy of subsequent links, to non-symmetric nomologically simultaneous relations of the envisaged kind. The familiar point is something for which Kim allowed in earlier work (Kim 1993 [1989]: 252). I can see no reason for rejecting the generalization to this case.

The point just made is also compatible with, indeed supports, the following possibility that may be behind Kim's reasoning. Consider a world in which emergent property dualism is true. There will be a possible world in which the same pattern of narrowly physical property causes over time will be instantiated without the emergent properties being present. This doesn't threaten the efficacy of emergent properties in the first world, however. The world that does not need emergent properties for the causal relations to hold will be a world with different laws to the laws that hold if emergent property dualism is the case. What proves to be insufficient in some worlds in virtue of the laws may be sufficient in others. Nor does this point rest upon rejecting the idea that a powers ontology is true (the laws being fixed simply by the kinds of properties that are instantiated), given the points made in section 1.

Emergent causation in the second sense identified—that involving either emergent or submergent consequences—does not seem to require the existence of emergent properties as causes. Indeed, given that emergent properties, themselves, are emergent consequences, it appears that there must be at least some emergent causation in which there are no emergent causes.

The suggested characterization of emergent causation in the second sense is a relatively straightforward application of the account of the distinction between the broadly physical and the emergent outlined in section 1.

xAy *strongly supervenes* on xBy just in case, necessarily, for each x and each causal relation in A , if xFy , then there is a pattern of (micro-)causal relations G such that xGy and, necessarily, if any x, y stand in G , then they stand in F i.e. $\Box (x)(y)(xFy \ \& \ F \ e \ A \ \Box (G) (G \ e \ B \ \& \ xGy \ \& \ \Box (u)(v)(uGv \ \Box \ xFy)))$.

A-relations are complex characterizations of exactly the way in which two things are causally related and so two things may stand in many such causal relations. The B-relations are characterizations of patterns of causal relations between x and y 's proper parts, for example, if x and y are property instances, then they hold between parts of their minimal supervenience-bases. Thus x stands in G to y if (say) x has parts $p_{x1}, p_{x2}, p_{x3} \dots p_{xn}$ which are causes of the respective parts of y , $p_{y1}, p_{y2}, p_{y3} \dots p_{yn}$. Suppose that, in a wide range of cases, perhaps in all, if a part of the supervenience-base of x is a cause of p_{yi} , then x is a cause of p_{yi} (in virtue of that part). It would not follow that an emergent causal relation does not hold between x and y . There may be a causal relation between the parts and also an emergent causal relation between x and y taken as a whole. Hence the importance of not thinking of the A-relations as simply whether or not a causal relation holds.

The second modal operator is once more to be understood as that of metaphysical necessity. If a case of non-emergent causation is just an arrangement of causal relations between instances of narrowly physical properties, then metaphysically necessarily, with this arrangement, non-emergent causation will be present. On the assumption that part of the supervenience-base for the existence of causal relations are the laws which cover them (when they do), then laws implicated in the existence of causal relations between instances of narrowly physical properties will be part of the legitimate supervenience-base of causal relations between instances of broadly physical properties. If these laws imply that instances of broadly physical properties will have causal relations which are not simply an arrangement of the causal relations of instances of narrowly physical properties, then, once more, these causal relations will hold in all possible worlds in which the supervenience-base is realized. Emergent causation will occur when neither of these two circumstances is met, most specifically, when the laws which explain how an arrangement of narrowly physical properties has certain causal consequences are not those which are implicated in the causal relations between instances of narrowly physical property instances alone. Since these laws may fail

to hold while the laws concerning narrowly physical properties still hold, the second modal operator would then only be that of nomological necessity.

One concern about this characterization of emergent causation is that it is too strong. There may be emergent causation if laws covering narrowly physical properties have distinctive consequences when applied to their arrangements. There are two possibilities here. Either the laws that cover the instances of the narrowly physical properties of the arrangement have distinctive consequences for the arrangement or laws that cover other instances of narrowly physical properties have distinctive consequences for an arrangement which does not involve these other instances. The latter would happen if properties of the arrangement were not shared by the property instances arranged. To give a toy illustration, laws covering square things may apply to square particles and square arrangements of round particles.

Either way, this would make emergent causation extremely ubiquitous and weak. Even if the law covering narrowly physical properties is additive, the result may yield distinctive causal consequences—for example, the smashing of a window by a baseball—which cannot be described simply in terms of the addition of the various causal contributions of the instances of narrowly physical properties but is partly the result of the nature of the thing to which these contributions are applied. Perhaps such cases might be excluded—and the notion of emergence thereby become stronger—if it is not so much a matter of whether there are distinct causal consequences as a result of the nature of the interacted upon but rather simply whether the laws are additive or not. Additive laws—such as additive force laws—are of particular significance if there are unit values for the determinable properties the law concerns. In that case, a law concerning the determinable properties can be reduced to laws about determinate properties with these unit values. So there is certainly a type of emergence to recognize here. It is a version of the weak emergence characterized in the previous section applied to causal relations. Carl Gillett has argued that this was the position of the pre-eminent British Emergentist, Samuel Alexander (Gillett 2006: 275–6, 285–6, and references to Alexander’s work therein; Alexander 1920). By contrast, Brian McLaughlin holds that Samuel Alexander, and others, supposed that there were fundamental forces which were only exerted by types of configurations of particles (McLaughlin 1992: 52, 66–7).

A second concern about emergent causation (understood in the way I have recommended) is that any causal powers which are putatively possessed by the emergent property seem legitimately attributed to narrowly physical properties instead (Ginet suggested this, as reported by O’Connor (1994: 89–99)). Suppose that $A(p_i, p_{i+1}, p_{i+2} \dots)$ is sufficient for the instantiation of a particular emergent property, E_1 and that, in circumstances C , E_1 putatively causes E_2 . Then p_i has the following power, in circumstances $A(\dots, p_{i+1}, p_{i+2} \dots)$ plus C , p_i is a cause of E_2 . In which case, any candidate case of emergent causation would turn out to involve causal relations between instances of narrowly physical

properties alone, so satisfying my analysis of non-emergent causation (see also Shoemaker 2002).

Considerations of theoretical simplicity are unlikely to speak in favour of postulating an emergent property. Although we would need only to attribute the disposition to instantiate E_1 in circumstances C to p_i , the number of dispositions we are attributing overall would be greater. There would be all the dispositions attributed as a result of the instantiation of E_1 together with p_i 's disposition to instantiate E_1 (as O'Connor and Wong acknowledge (2005: 672)).³

Special arrangements of instances of narrowly physical properties, those with distinct causal powers, are emergent phenomena. The point is not that, if we recognize special arrangements of such properties, we can avoid recognizing an emergent phenomenon (see O'Connor 2000a: 113–14). The issue is rather whether we need to recognize emergent properties in addition to emergent causation. There does not seem to be any reason to do so. Of course, if a causal theory of properties is true, then the distinct causal powers attributed to the arrangement of instances of narrowly physical properties, rather than the instances of narrowly physical properties themselves, will be sufficient to establish the existence of an emergent property. Nevertheless, it is important to recognize that this additional premise is required.

The correct response to the second concern is that emergent causation is present, not because causal relations between instances of narrowly physical properties fail to necessitate it, but because the laws envisaged attribute fundamental forces to narrowly physical properties when they occur as part of a certain kind of arrangement or, in some way, seem isolated from the other laws that hold. Such laws should not figure as part of the supervenience-base in the account given earlier. This is more important than to which property the causal power should be attributed. Nevertheless, there can be reason to postulate the existence of emergent property, either because of the pattern of causal relations we observe or because we have other, non-causal reasons for believing in its existence, for example to characterize the nature of our experience. In the last section of this chapter, I summarize the various kinds of emergence I have distinguished and relate them to two possible cases of emergent phenomena.

5. CONCLUDING REMARKS AND POSSIBLE CASES OF EMERGENCE

I have argued that strongly emergent properties are to be distinguished from broadly physical properties by the fact that the strongly emergent properties

³ Unfortunately, for reasons of space, I cannot discuss O'Connor and Wong's argument that without postulating emergent properties we would be left with action at a distance (see O'Connor and Wong 2005: 672–3).

are only related to their supervenience-base by nomological necessity. Weakly emergent properties are a subcategory of broadly physical properties in which the properties are in some way hard to infer from arrangements of narrowly physical properties. Causation by strongly emergent properties is a species of emergent causation. Nevertheless, broadly physical properties may also stand in emergent causal relations. These also come in strong or weak versions. According to the strong version, emergent causal relations are related only by nomological necessity to patterns of causal relations between instances of narrowly physical properties. Weak emergent causation just insists that narrowly physical property instances should have different powers when taken together in a particular arrangement than they would taken in isolation. Physical laws will reflect this.

Quantum mechanics has been thought to support emergent physicalism, that is, narrowly physical properties which are emergent from other narrowly physical properties. Suppose that emergent physicalism is true. Is there any reason to be an emergent non-physicalist due to mental phenomena?

Phenomenal consciousness does not seem to be captured by recognizing emergent causation by non-emergent properties. We have little evidence for the existence of emergent causation as a result of phenomenal consciousness unless we take seriously the idea that phenomenal consciousness can only be captured by non-physical phenomenal properties. Intercranial causal relationships seem to be nothing out of the ordinary. Indeed, that's partly why the causal completeness of the physical realm seems so plausible. Of course, if there are emergent phenomenal properties, then there will be emergent causation in both senses. For that reason, from the point of view of the philosophy of mind, emergent physicalism seems to me to be an unattractive doctrine. It postulates emergent causation for no reason and yet denies that phenomenal properties are broadly non-physical. This reservation does not apply to Gillett's emergent physicalism only involving weakly emergent causation (Gillett 2003, 2006).

Should we be forced to accept the existence of non-physical emergent phenomenal properties, this won't just be because it is difficult to attribute a predicate to anything other than a system of entities or because it is difficult or impossible to attribute a predicate on the basis of the knowledge of boundary conditions plus dynamic laws (see Heard (2006: 58–61), who attributes these commitments to the emergentist). If it is simply difficult but not impossible to attribute a predicate, or only epistemically impossible for McGinn-style closure reasons, then there are no emergent properties in the sense we have identified. No sensible emergentist is going to accept that there is a straightforward inference from the fact that something can only be predicated of a whole system, and that the development of a system can be understood in terms of laws that make no reference to the activity of elements, to the claim that there exist emergent properties. Rather, this proposal will be assessed in the light of competing considerations of which the observation just mentioned is only one. Deciding that there are emergent non-physical phenomenal properties will be the complex result of balancing causal

arguments for physicalism against different possible accounts of the explanatory gap and our introspective experience. Thus O'Connor emphasizes the apparent non-structurability and subjectivity of phenomenal properties, and Kim that they apparently cannot be functionalized (O'Connor 2000a: 116–17; Kim 2005).

Libertarian accounts of free will provide another potential case of emergence. According to some, agents have a property in virtue of the fact that they are able to cause events to happen in their bodies (see, e.g., O'Connor 2000a: 121–3). A certain way of understanding libertarian free will may seem to threaten the claim that emergent properties even N-supervene upon narrowly physical properties. Suppose that $A_1(p_1, p_2, p_3, \dots)$ indeterministically causes $A_2(p_{100}, p_{101}, p_{102}, \dots)$ and in one subject, when in $A_1(p_1, p_2, p_3, \dots)$, the subject considers whether to imagine sensuously purple, decides to and, when in $A_2(p_{100}, p_{101}, p_{102}, \dots)$, he or she does. In another subject, his or her decision may go the other way. If you insist that $A_2(p_{100}, p_{101}, p_{102}, \dots)$ will still be caused by $A_1(p_1, p_2, p_3, \dots)$, then this can't be even the N-supervenience-base for imagining sensuously purple. The argument can presumably generalize (see O'Connor 2000b: 111–12).

Two points should be recognized. The first is that this way of understanding libertarian free will is not required. We can keep the connection between $A_2(p_{100}, p_{101}, p_{102}, \dots)$ and imagining sensuously purple so long as the decision to imagine another colour instead indeterministically causes a different arrangement of narrowly physical properties $A_3(p_{1000}, p_{1001}, p_{1002}, \dots)$. There is no reason to suppose that the physico-psycho law is indeterministic. Indeed, there is no need to suppose that libertarian free will involves additional emergent properties. We can just suppose that there is emergent causation. If the world is indeterministic, there will be a range of possible events which might occur in agents' brains. Which one does on any occasion is not just by chance nor by law, but directly influenced by an agent. Nor need such a picture involve a revision to the laws of physics since the range of alternatives are set by them. All we would have is one unlikely sequence rather than another actualized by the agent. If the causation in question is a case of agent causation, then it will be emergent not because of the application of some new law but because of the activity of an agent.

The second point is that, if there is interlevel indeterminism (and why not), then the supervenience claim should be understood correspondingly. Thus N-supervenience will postulate an indeterministic but still nomic relationship between narrowly physical properties and mental properties. It will no longer be appropriate to say that if two subjects are identical in their narrowly physical properties, then they should be identical in their mental properties. Rather, the claim will be that the probabilities of various mental properties will be the same.

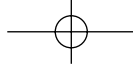
However, as I have already indicated, this dimension of the position is not mandatory.

Thus, we see that the distinction between emergent properties and emergent causation drawn in this chapter enables us to differentiate between emergent approaches to the mind that are a challenge to physicalism and those which are not.

REFERENCES

- Alexander, S. 1920. *Space, Time and Deity: The Gifford Lectures 1916–1918*, 2 vols. London: Macmillan.
- Bedau, M. A. 1997. 'Weak Emergence'. In J. E. Tomberlin (ed.), *Philosophical Perspectives* 11: 375–99.
- Bennett, K. 2003. 'Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It'. *Noûs* 37.3: 471–97.
- Chalmers, D. 2006. 'Strong and Weak Emergence'. In P. Clayton and P. Davies (eds), *The Re-Emergence of Emergence*. Oxford: Oxford University Press.
- Davidson D. 1980 [1963], 'Actions, Reasons and Causes'. *Journal of Philosophy* 60, reprinted in his *Essays on Actions and Events* (Oxford, Oxford University Press), pp. 3–19.
- Ganeri, J., Noordhof, P. and Ramachandran, M. 1996. 'Counterfactuals and Preemptive Causation'. *Analysis* 56: 216–25.
- 1998. 'For a (revised) PCA Analysis' *Analysis* 58.1: 45–7.
- Gibbons, J. 2006. 'Mental Causation without Downward Causation'. *The Philosophical Review*, 115.1: 79–103.
- Gillett, C. 2003. 'Non-Reductive Realization and Non-Reductive Identity'. In S. Walter and H.-D. Heckmann (eds), *Physicalism and Mental Causation*. Charlottesville, VA: Imprint Academic, 31–57.
- 2006. 'Samuel Alexander's Emergentism: Or, Higher Order Causation for Physicalists'. *Synthese* 153: 261–96.
- and Rives, B. 2005. 'The Non-Existence of Determinables: Or, a World of Absolute Determinates as Default Hypothesis' *Noûs* 39.3: 483–504.
- Heard, D. 2006. 'A New Problem for Ontological Emergence'. *Philosophical Quarterly* 56.222: 55–62.
- Heller, M. 1998. 'Property Counterparts in Ersatz Worlds'. *Journal of Philosophy* 95.6: 293–316.
- Hüttemann, A. and Papineau, D. 2005. 'Physicalism Decomposed'. *Analysis* 65: 33–9.
- Jackson, F. 1998. *From Metaphysics to Ethics*. Oxford: Oxford University Press.
- Kim, J. 1993 [1984]. 'Concepts of Supervenience'. In idem, *Supervenience and Mind*. Cambridge: Cambridge University Press, 53–78.
- 1993 [1989], 'Mechanism, Purpose or Explanatory Exclusion'. In *Supervenience and Mind*. Cambridge: Cambridge University Press, 237–264.

- Kim, J. 1998. *Mind in a Physical World*. Cambridge, MA: The MIT Press.
- 1999a. 'Supervenient Properties and Micro-Based Properties: A Reply to Noordhof'. *Proceedings of the Aristotelian Society* 99: 115–18.
- 1999b. 'Making Sense of Emergence'. *Philosophical Studies* 95.1–2: 3–36.
- 2005. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press.
- Lewis, D. 1986a [1979], 'Counterfactual Dependence and Time's Arrow'. *Philosophical Papers*, vol. 2. Oxford: Oxford University Press, 32–52.
- 1986b. *On the Plurality of Worlds*. Oxford: Basil Blackwell.
- McGinn, C. 1991 [1989], 'Can we solve the Mind-Body Problem?' *Mind* 98.391: 349–66, repr. in his *The Problem of Consciousness* (Oxford: Basil Blackwell, 1991), 1–22.
- McLaughlin, B. 1992. 'The Rise and Fall of British Emergentism'. In A. Beckermann, H. Flohr, and J. Kim (eds), *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. New York: Walter de Gruyter, 49–93.
- Noordhof, P. 1997. 'Making the Change: the Functionalist's Way'. *British Journal for the Philosophy of Science* 48: 233–50.
- 1999a. 'Causation by Content?'. *Mind and Language* 14.3: 291–320.
- 1999b. 'Micro-based Properties and the Supervenience Argument: a Response to Kim'. *Proceedings of the Aristotelian Society* 99: 109–14.
- 1999c. 'Probabilistic Causation, Preemption and Counterfactuals'. *Mind* 108.429: 95–125.
- 2003. 'Not Old . . . But Not That New Either: Explicability, Emergence and the Characterization of Materialism'. In S. Walter and H.-D. Heckmann (eds), *Physicalism and Mental Causation*. Charlottesville, VA: Imprint Academic, 85–108.
- 2006. 'Environment-Dependent Content and the Virtues of Causal Explanation'. *Synthese* 149: 441–575.
- O'Connor, T. 1994. 'Emergent Properties'. *American Philosophical Quarterly* 31.2: 91–104.
- 2000a. *Persons and Causes*. Oxford: Oxford University Press.
- 2000b. 'Causality, Mind and Free Will'. In J. E. Tomberlin (ed.), *Philosophical Perspectives* 14. Boston MA and Oxford: Blackwell, 105–17.
- and Wong, Hong Yu. 2005. 'The Metaphysics of Emergence'. *Notus* 39.4: 658–78.
- Papineau, David (forthcoming). 'Must a Physicalist be a Microphysicalist?'
- Ramachandran, M. 1997. 'A Counterfactual Analysis of Causation'. *Mind* 106: 263–77.
- Shoemaker, S. 1980, 'Causality and Properties'. In P. Van Inwagen (ed.), *Time and Change* (Dordrecht: D. Reidel) and repr. in his *Identity, Cause and Mind* (Cambridge: Cambridge University Press, 1984), 206–33.
- 1981, 'Some Varieties of Functionalism'. *Philosophical Topics* 12.1: 83–118, repr. in his (2003), *Identity, Cause and Mind: Expanded Condition* (Oxford: Oxford University Press), 261–86.
- 2002. 'Kim on Emergence'. *Philosophical Studies* 108: 53–63.
- Silberstein, M. 2006. 'In Defence of Ontological Emergence and Mental Causation'. In Philip Clayton and Paul Davies (eds), *The Re-Emergence of Emergence*. Oxford: Oxford University Press, 203–26.
- Sober, E. 1999. 'Physicalism from a Probabilistic Point of View'. *Philosophical Studies* 95.1–2: 135–74.



- Van Cleve, J. 1990. 'Mind-Dust or Magic? Panpsychism Versus Emergence'. In J. E. Tomberlin (ed.), *Philosophical Perspectives*, vol. 4: *Action Theory and Philosophy of Mind*. Atascadero, CA: Ridgeview, 215–26.
- Wilson, J. 2005. 'Supervenience-based Formulations of Physicalism'. *Noûs* 39.3: 426–59.
- Yablo, S. 1992. 'Mental Causation'. *The Philosophical Review* 101.2: 245–80.

