European Journal of Philosophy 2008, 17:3, pp. 347-362

Levels, orders and the causal status of mental properties

penultimate draft

Simone Gozzano

Abstract

In recent years Jaegwon Kim has offered an argument – the 'supervenience argument' – to show that supervenient mental properties, construed as second-order properties distinct from their first-order realizers, do not have causal powers of their own. In response, several philosophers have argued that if Kim's argument is sound, it generalizes in such a way as to condemn to causal impotency all properties above the level of basic physics. This paper discusses Kim's supervenience argument in the context of his reply to this so-called 'generalization argument'. In particular, the paper focuses on the level/order distinction, to which Kim appeals in his reply to the generalization argument, and on the relation between this distinction and two varieties of functionalism, 'realizer' *vs.* 'role' functionalism. The author argues that a proper analysis of the notions of *levels* and *orders* undermines Kim's response to the generalization argument, and suggests that Kim's reductionist strategy for vindicating the causal powers of mental properties is better served if mental properties are construed as first-order properties, as realizer-functionalism recommends.

1. Introduction

Jaegwon Kim's 'supervenience argument', as formulated in his (1998) *Mind in a Physical World* and recently developed and clarified in more recent works (2003, 2005), is meant to show that mental properties, construed as second-order properties distinct from their first-order realizers, do not have causal powers over and above those of their realizers. Against Kim's argument, some have proposed the 'generalization argument', whose aim is to show that Kim's argument, if sound, would condemn to causal impotency all kinds of properties except, perhaps, those of basic physics. In order to defend his position against such an unappealing consequence, Kim proposes a distinction between *levels* and *orders*, arguing that the supervenience argument applies just to second-*order* properties (as mental properties are, on his account), and not to

'special science' properties in general, these being located at different *levels*. The purpose of this paper is to evaluate Kim's response to the generalization argument in the context of his account of the relation between functionalism and reductionism, and in the light of his view that the world can be understood as organized in a hierarchy of levels.

After summarizing the dialectics of the relationship between the supervenience and the generalization arguments, I will analyze the level/order distinction and its relation to two varieties of functionalism, 'realizer' *vs.* 'role' functionalism.[1] I will attempt to show that Kim's way of interpreting the level/order distinction is not consistent with his overall position, or else fails to apply to the functionalist interpretation of mental properties. I will further argue that a proper analysis of the notions of 'level' and 'order' undermines Kim's supervenience argument, by showing that mental properties are better construed as *first-order properties*, as realizer-functionalism recommends. Such an analysis, moreover, permits a better solution to the problem of the causal status of mental properties – a solution that is consistent with Kim's ultimate reductionist aims.

2. The supervenience argument and the generalization argument

For some authors (e.g., Tyler Burge 1993 and Lynne Rudder Baker 1993), the issue of mental causation is essentially an epistemological one: it is not *whether there is* mental causation, but *how we are able to know*, or *justify* our claim to know, that there is. For Kim, instead, the issue is essentially a metaphysical one: granted the reality of mental causation, the problem is to *explain how it is possible* – that is, '*how to make our metaphysics consistent with mental causation'* (Kim 1998: 62). The supervenience argument is intended to articulate the nature of this metaphysical problem in the context of his critique of nonreductive physicalism.

Kim (2005) thinks that every non-reductive physicalist should accept the following metaphysical theses:

Supervenience: mental properties strongly supervene on physical properties, that is, if any system *S* instantiates a mental property *M* at *t*, then it necessarily instantiates some physical property *P* at *t* and, necessarily, any *S* instantiating *P* at *t* thereby instantiates *M* at *t*;

Closure: if a physical event has a cause at *t*, it has a physical cause at *t*.

Irreducibility: mental properties are irreducible to, and not identical with, physical properties.

Given these three theses, the supervenience argument proceeds as follows: Assume that (an instance of) a mental property *M* causes another mental property *M** to be instantiated. By *Supervenience* both *M* and *M** have subvenient physical bases which realize them; call these *P* and *P**, respectively. Assuming, as seems reasonable, that something cannot cause a supervenient property to be instantiated without causing its base, it seems that M caused M* by causing M*'s supervenient base, *P**. However, by *Closure*, P* must have a physical cause, and since P is sufficient for M and M is (*ex hypothesis*) sufficient for P*, it is reasonable to suppose that P is a sufficient cause of P*. But then we have both M causing P* and P causing P*; and since, by *Irreducibility*, M ≠ P, we are led to the conclusion that *M** has two distinct sufficient causes: M and P. At this point Kim supplements the above three theses with a fourth:

Exclusion: no event can have more than one sufficient cause, unless it is a genuine case of overdetermination.

Since for Kim it is not plausible to suppose that here we have a case of overdetermination,[2] *Exclusion* forces us to choose between M and P as the real cause of P*, and, by *Closure*, our choice must be P*. The causal link between P and P* is thus sufficient for the occurrence of M* , and no causal work is left for M to do. Since the argument applies to all mental properties, Kim's conclusion is that nonreductive physicalism is committed to the causal inertness of mental properties, i.e., epiphenomenalism.

Why does Kim thinks that this argument applies specifically to mental properties and not, or not necessarily, to other 'special science' properties?

In accordance with classic functionalism[3], Kim construes mental properties as second-order properties, which he defines as follows: '*F* is a *second-order property* over set **B** of base (or first-order) properties iff *F* is the property of having some property *P* in **B** such that *D(P)*, where *D* specifies a condition on members of **B**' (Kim 1998: 20). Consider Kim's own example: the property of having a *primary colour* is the property defined as the property of having a *colour* which satisfies some further condition – the condition that if mixed with other appropriate colours it produces the entire visible

spectrum[4]. As will be noticed, the logical status of second-order properties, as defined by Kim, is that of a proper subset of the set to which the first-order properties belong. That is to say, the domain of the property of having a *primary colour* is properly nested in the (nonempty) domain of the property of having a *colour*, hence, unsurprisingly, the former set supervenes on the latter. And so too for mental properties construed as second-order properties.

The following consideration also allows us to see how the conclusion of the supervenience argument – that mental properties have no causal powers of their own – is supported by Kim's construal of mental properties as second-order properties. Having a primary colour is equivalent to having either red or blue or green. Having red is a first-order property; so, if something has a primary colour by having red (or being red) then the second-order property is realized in that thing through its possessing that specific first-order property, namely red. Red, however, does not acquire new causal powers in virtue of its being a primary colour. In general, we can say, the causal powers of any second-order property on each occasion of its instantiation turn out to be the same as the causal powers of its realizer on that occasion. It follows that second-order properties have no causal powers of their own, over and above those of their realizers.[5]

Several authors[6] have argued that Kim's supervenience argument can be generalized so as to apply to any kind of supervenient property. This so-called 'generalization argument' aims to show that, if the supervenience argument is sound, it condemns to causal impotency *all* supervenient properties. If, in general, special-science properties supervene on basic physical properties, the supervenience argument should apply to the properties of all sciences with the exclusion of basic physics, the only level supervening on no other, where all the real causal powers are ultimately to be found. But, on the face of it, there seems to be no problem with admitting chemical or biological causation, and so there seems to be genuine causation above the level of basic physics. In the absence of an argument showing why psychological properties should be treated any differently from chemical or biological properties, there seems to be no reason to worry about the causal status of mental properties.[7] And so, it has been claimed, something must be wrong with the supervenience argument.

Kim, however, denies that his supervenience argument does so generalize, and to show this he appeals to the distinction between *levels* and *orders*. The generalization argument, Kim argues, mistakenly assumes that the supervenience relation, which in the context of the supervenience argument applies to second-order mental properties and

their first-order realizers, also holds of properties ordered by the macro-micro relation, which is a relation between *levels*. The illusion that the supervenience relation applies to the macro-micro hierarchy is generated by the fact that in many interesting cases the first-order realizers of second-order macro-properties are *micro-based* (or *microstructural*) properties – properties that macro-objects have in virtue of their mereological composition.[8] But, Kim stresses, these micro-*based* properties are not *microproperties*: rather, they are themselves *macro*properties occurring at the same level as the macro-properties they realize. Since, for Kim, the supervenience relation (as defined above) is exclusively an *intralevel* relation, properties occurring at different levels along the macro-micro hierarchy are not related by supervenience, and so the supervenience argument does not generalize to them.

In order to assess the soundness of Kim's reply, we need to take a closer look at the levels/orders distinction and see how it applies to the specific case of mental properties.

3. The layered model

The general intuition behind the layered model is that reality is made up of many levels, each with its own properties and relations. Every layer is connected in some specific way to the layers below and above it. On most accounts, the hierarchy has both an ontological and an epistemological dimension. According to Elliot Sober, for example, '[t]he domains of higher-level sciences are subsets of the domains of lower-level sciences … In addition, since the domains are (properly) nested, there will be phenomena that lower-level sciences can explain but that higher-level sciences cannot' (Sober 1999: 543). But how are levels to be individuated?

According to Kim levels are individuated by reference to a hierarchy of *objects* construed as entities or systems ordered by a mereological relation (Kim 1998: 82ff). The basic constraint on assigning properties to levels is that if two properties can be attributed to the very same entity, they pertain to the same level. For instance, metabolic and developmental properties that are attributed to biological objects belong to the same level. Properties present at the level of biological objects are not present at lower levels - for instance, they are not present at the level of chemical or physical objects. Properties present at the level of chemical objects may or may not be present at upper levels depending on whether they can be ascribed to, for instance, biological objects (and not just to their parts).

This view, clearly, departs from the usual view according to which levels are the metaphysical counterparts of the distinctive subject matter of the various empirical sciences, such as physics, chemistry, biology, etc. Thus my weighing 65 kg, my believing that *p*, and my being of Italian nationality, which the classical view would assign to different levels (physical, psychological, political, respectively) are for Kim all at the same level, being all properties of *me*. Clearly, the classical view would not recognize a level that simultaneously comprises physical, psychological and political properties. On Kim's picture, the supervenience relation just applies to properties belonging to the same level, for instance it holds between *my having* certain psychological properties and *my having* certain neurological properties.[9] However, there are properties that apply to mereological *parts* of myself, but not to me as a whole (e.g. C-fiber firing is predicable of *my c-fibers*, not of *me*). This raises a question about the relation between properties that apply to me as a whole and properties that apply to my proper parts, given that the respective sets of properties belong to different levels. Since for Kim supervenience does not characterize the relation between levels but that between orders within the same level, the question arises how Kim's layered model can support a doctrine of *mereological supervenience*, 'the doctrine that properties of wholes are fixed by properties and relations that characterize their parts' (Kim 1998: 18). This question arises because the properties and relations that characterise the parts exist at a lower level than the properties and relations that characterise the whole object, and for Kim supervenience (as defined above) requires the coinstantiation of the supervenient and subvenient properties in the *same object*, and thus at the *same level*. This is a crucial metaphysical issue, as Kim recognizes – an issue with consequences for reductionism and emergentism. As we shall see, it is an issue that has some problematic consequences for Kim's response to the generalization argument.

4. Properties: orders and levels

What is the relation between levels and orders? As already indicated, for Kim second-order properties occur at the same level as their first-order realizers. More specifically, Kim holds, '[t]he series created by the second-order/realizer relation does not track the ordered series of micro-macro levels; it stays entirely within a single level in the micro-macro hierarchy' (1998: 82), and the reason, he adds, is that '*a second-order property and its realizers are at the same level in the micro-macro hierarchy; they are properties of the very same objects*' (ibid., Kim's italics).[10] This, of course, is

supposed to apply to mental properties and their physical realizers as well, in so far as the former are construed as second-order and the latter as first-order properties. Let me call this the *same-level* view of the relation between (second-order) mental properties and their (first-order) physical realizers.

However, Kim also says the following about the relation between mental and physical properties: 'When mental properties are to be generated out of **B** as second-order properties (where **B** is the set of first order base properties), we must *of course* take **B** to consist of *nonmental properties* (including physico-chemical, biological, and behavioural properties)' (ibid.: 20, italics added). It seems reasonable to suppose that these *nonmental* properties should be attributable to some subpersonal system, that is, to parts of a person (or an animal or artifact) taken as a system, while the mental properties themselves are attributable only to the person or (whole) system. Being attributable to different objects in the mereological hierarchy, these mental and nonmental properties lie at different levels. Consider, for example, my feeling pain and the firing of my C-fibers. The former is surely a property of mine; what about the latter? *My having* my C-fibers firing is a (micro-based) property of mine, but *the firing* of my C-fibers is surely not a property of mine but of my C-fibers. Hence the two properties – my being in pain and the firing of my C-fibers (which one would expect to be part of the supervenience base of my being in pain) – occur at different levels on Kim's sense of levels. So it is reasonable to take the above condition on the 'generation' of mental properties as indicating that mental properties can occur at a different level from that of their realizers. Let me call this the *different-level* view of the relation between mental properties and their physical realizers. Which view does Kim actually endorse: the same-level or the different-level view? Let us take a closer look at the different-level view first – the one that Kim would likely be more suspicious of.

4.1. Different-level

Let us suppose that mental properties and their subvenient physical realizers are at different levels. If so, it is implicitly admitted that there is a level that mental properties belong to. What level is this? A plausible answer is: the psychological – that is, personal - level, the one at which these properties are attributed to persons. As Kim says: 'psychology is a special science located at one of these levels, toward the higher end, in this multilayered system, and mentality is a distinctive set of properties that make its appearance at this level' (ibid.: 79). In this passage Kim seems to agree with the

classical view concerning the individuation of levels, according to which levels are individuated by distinctive sets of properties. Indeed, if levels were individuated with reference to objects, then, if I were to lose all my psychological properties, many of my other properties (e.g. biological and physical properties) would change, but their *level* would not, since the object to which they are ascribed remains the same.

Now, if there is a distinctive set of properties that make up a level, one may wonder whether this set is constituted by first- or by second-order properties. Consider the psychological level. If this level were constituted just by second-order properties (as Kim takes mental properties to be) then, given the conclusion of Kim's supervenient argument, this level of properties would turn out to be causally inert. Such a level of properties could not be invoked in causal explanations but would at best have an instrumental use. On the other hand, if the level were constituted by first-order properties, their causal powers would not be threatened by the 'supervenience argument', since this only calls into question the causal efficacy of *second*-order properties, and so mental properties would be as efficacious as any other kind of first-order property.

Suppose we uphold both the soundness of the supervenience argument and the (role-) functionalist interpretation of mental properties as second-order properties. It follows that the psychological level, classically understood as a distinctive level of *properties*, is causally inert. To avoid this consequence Kim might stick to his original way of individuating levels in terms of a hierarchy of objects, not of properties. In this case, however, consider again my feeling pain and the firing of my C-fibers. As remarked earlier, the second-order property of feeling pain is a property of mine, and so too is the property of my having my C-fibers firing. But the firing of my C-fibers is a property of my C-fibers, not of me, and so this property and the former two properties are at different levels on Kim's own conception of levels. But, now, Kim takes the supervenience relation to be an *intra*level, not an *inter*level relation, and so it follows that neither my property of feeling pain nor my (micro-*based*) property of having my C-fibers firing supervene on the firing of my C-fibers – the latter being a (lower-level) property of my C-fibers, not of me. This entails that there can be variation in my feeling pain or in my having my C-fibers fire without a corresponding variation in the firing of my C-fibers, even when these are properly 'wired' in my organism – surely an implausible result!

The problem, to put it differently, is that if second-order mental properties are 'generated' out of the set of their first-order realizers, then they should supervene on their realizers because, once you have fixed the set of the realizers you have fixed the set of second order properties generated out of it. It makes no difference whether the first order realizers are at the same level as, or at a lower level than, the second-order properties generated out of them. If mental properties are regarded as second-order properties with physical properties as their first-order realizers, and, at the same time, the relation between mental and physical properties is one of supervenience, then the relation between first and second-order properties should be of supervenience too. So, if Kim subscribed to the 'different-level view' of the relation between first- and second-order properties while denying interlevel supervenience, he could not coherently uphold mind-body supervenience.

On the other hand, if Kim allowed supervenience as an interlevel relation, then the motivation for the distinction between orders and levels would be lost, for then the series generated by the second-order/realizer relation would, after all, *track* the ordered series of macro-micro levels. And this, in turn, would amount to endorsing the generalization argument, thereby giving credit to the view that the causal status of supervenient mental properties is no different than the causal status of any other 'special-science' supervenient property.

Carl Gillett (2002) has insisted that if interlevel supervenience is denied, then we cannot take advantage of the many explanations that empirical sciences have provided on how the causal powers of the microcomponents. For instance, to use Gillett's example, cut diamonds scratch glass because their carbon atoms are bonded and alligned in a way to form lattice-like structures that maintain their respective positions even under high temperatures and forces. The properties of being so bonded and alligend are of the carbon atoms, not of the diamond; the power of scratching glass is of the cut diamond, not of the atoms: nevertheless, we explain the causal powers of the latter in terms of the causal powers of the former, and this assumes interlevel supervenience.

Bontly (2002) has also argued that there is no good reason why Kim should impose a 'coinstantiation principle' on supervenience, according to which if x's having property F supervenes on y's having property G then x=y, thereby ruling out interlevel (macro-micro) supervenience. Supervenience, as such, need not require that the properties that co-vary belong to the same object. Indeed, Bontly points out that Kim

himself once provided an account of supervenience 'for multiple domains' that explicitly allows, under suitable (merelogical) constraints, for supervenience relations to hold between properties belonging to different domains of objects (Kim 1988). This broader notion of supervenience allows us to say not only that *my being* in pain supervenes on *my having* my C-fibers firing (an instance of intralevel supervenience), but also that my being in pain supervenes on *the firing of my C-fibers* (an instance of interlevel supervenience). If so, the exclusion problem does generalize, for the *Exclusion* principle does not require that the supervenient and the subvenient cause must be properties of the same object. If my having my C-fiber firing excludes my being in pain as a sufficient cause of my wincing, then so too does the firing of my C-fibers. And the same holds for any supervenient 'special science' property vis-à-vis the lower-level properties on which they happen to supervene.

Thus intralevel supervenience should not be thought as excluding mereological supervenience understood as an interlevel relation. It's true, as Kim point out, that 'macroproperties can, and in general do, have their own causal powers, powers that go beyond the causal powers of their microconstituents'. But when these microconstituents are in appropriate mereological configurations, their properties and relations do determine the causal powers of a system's macroproperties. It's also true, as Kim explains, that mereological supervenience requires that 'the fact that [system] s has [property] P … is fixed once the micro-constituents of s and the properties and relations characterizing these constituents are fixed', and that this means that 'the base property on which P supervenes is a micro-based property, the property of having such-and-such proper parts that have such-and-such properties and that are configured by such-and-such relations. This is a micro-based property of s, not a property that belongs to any of its proper parts' (1998, pp. 85-86). But this is consistent with this micro-based property of s being determined by, and thus supervening on, mereological configurations involving microproperties of proper parts of s.[11]

It should be noticed, finally, that even considering *intra*level and *inter*level supervenience as two completely independent relations, would not avoid the mentioned problem. In fact, it is possible to rehearse the previous argument, making the case for the generalization argument, within each single level inteded *a la* Kim. For instance, consider a magnetic field into a ferromagnetic element: it has properties such as the intensity. It is realized by atoms, whose spin properties of their electrons, nuclear and orbital *momenta*, constitute it. We may consider these atoms, and their properties, as the

subvenient *intra*level mereological base of the magnetic field. Nevertheless, the spin properties of the electrons, nuclei and orbits of the atoms are not the intensity property of the field. In order to secure the determination relation one has to allow that the properties of the field supervene on the properties of its constituents, letting the causal power to seep down from the whole to its constituent parts.

The *different-level* view, then, either – by denying interlevel supervenience – leads to a quite untenable view concerning the relation between the properties of myself and those of my parts, or – by affirming interlevel supervenience – allows the tracking of the level-hierarchy by the order-hierarchy, with the consequence that the supervenience argument generalizes, and if the argument is sound, all causal powers seep down to the lowest level of reality– the level of basic physics.[12] Clearly, then, Kim could not coherently hold the different-level view of the relation between supervenient mental properties and their subvenient physical realizers while pursuing his strategy for resisting the generalization argument. Let us turn, then, to the *same-level* view.

4.2. Same-level

Unlike the different-level view, the same-level view appears *prima facie* consistent with the layered model as conceived by Kim, given his construal of mental properties as second-order properties. On this view, second-order mental properties belong to the same level as their first-order realizers, since they are predicable of the same individual objects. Second-order mental properties are nested within levels, as properties defined over those properties ascribable to individuals. Once a level has been identified relative to a category of individuals, the first-order properties ascribable to individuals in that category are at the same level as the second-order properties ascribable to them. The level at which these two sets of properties pertain is, as indicated at the beginning of section 4.1, that of persons or, more generally, of individuals, because mental properties are typically predicated of persons or of individuals to whom beliefs, desires, emotions, and other states of consciousness are properly attributable. It makes little difference, for our purpose, whether such properties are attributable only to human beings or also to other kinds of animals, or artificial or alien creatures: what matters is that these properties play the appropriate causal/functional roles in the behaviour of the creatures to which they are attributed.

Now, a problem looming behind Kim's response to the generalization argument *vis a vis* his adoption of the same-level view emerges from considering the layered

model against the backdrop of functionalism. As already indicated, there are two distinct varieties of functionalism. According to the first, *realizer-functionalism*, mental properties are identical with those first-order states or properties that occupy the functional roles that individuate them (cf. Lewis 1966; 1972; Jackson et al 1982). The basic idea is that neurological properties, for those species that have them, just *are* mental properties if they play the relevant causal roles in psychological explanations. Mental properties, on this view, are (identical with) first-order properties;[13] e.g., believing is the state that fills the appropriate causal role for a given organism (e.g. of guiding its behaviour in some representationally constrained way.)

According to the second version of functionalism – *role-functionalism* – mental states or properties are second-order properties (Block 1980; Loar 1981): to have a mental property is to have the property *of being in some state* (or *of having some property*) that plays a given role. The realizers of mental properties are first-order properties, while the mental properties themselves are second-order properties.[14] Believing is second-order in the sense of being the property of being in some internal state satisfying the further condition of playing such-and-such causal role. Since the various roles can have quite different realizing conditions, role-functionalism clearly allows for multiple realizability[15].

So the question is: what kind of functionalism is Kim assuming in his response to the generalization problem? It is natural to answer that it is role-functionalism, the only of the two versions that holds that mental properties are second-order properties. Now, is Kim's view compatible with role-functionalism? Role-functionalism tells us that mental properties are distinct from their first-order realizers, though, as Kim has made clear (p. 82), both kinds of property are predicated of the same object and hence are at the same level in the micro-macro hierarchy. Thus, *John* has the property of being in pain and *John* has the distinct micro-*based* property of having his C-fibers firing – the latter being the first-order property realizing John's second-order property of being in pain. But, obviously, John could not have the property of *having his C-fibers firing*, and thus of being in pain, unless *his C-fibers are firing*, that is, unless a proper (mereological) *component* of John has a certain micro-property; and that component's property, by Kim's own account, is at a lower-*level* along the micro-macro hierarchy than the first-order property that realizes John's property of being in pain.[16] Indeed, it is because that first-order property of John *supervenes on*, or is *determined by*, a configuration of lower-level properties that include the firing of his C-fibers, that John's

first-order property is able to play the appropriate pain-role, and thus realize John's property of being in pain. Unless supervenience applied across levels, and not only within levels, we could never determine that, or why, a second-order property has the first-order realizer it has, namely, one that plays the pain role for John. In other words, unless the firing of C-fibers (a lower-level, microproperty of the C-fibers) in creatures like John were a sufficient supervenience base for such creatures' being in pain, it would remain obscure why John's having his C-fibers firing (a higher-level, micro-based property of John) should realize his being in pain. By disallowing interlevel supervenience, Kim in effect makes it impossible to determine what same-level, first-order property plays the role that role-functionalism requires it to play. Hence espousing role-functionalism cannot serve Kim's strategy for coping with the generalization problem, since role-functionalism itself presupposes interlevel supervenience, which opens the way to the generalization argument.

Nor can realizer-functionalism be the version of functionalism that Kim can coherently embrace in his response to the generalization problem, for realizer-functionalism identifies mental properties with the first-order properties that play the requisite roles, whereas Kim's response to the generalization problem, as we have seen, assumes that mental properties are second-order properties. The upshot is that Kim's reply to the generalization argument is inconsistent with both versions of functionalism. Whether it might be consistent with some other version of functionalism[17] is a question that we cannot take up here.

5. The causal powers of mental properties

The discussion so far leaves us without a clear position concerning the causal status of mental properties. We should note that Kim generally endorses the view, promoted by Shoemaker[18], that 'real' properties in general are essentially individuated by their causal powers (or, if you will, by the causal powers they confer on the things that have them). As Shoemaker puts it, '..if under all possible circumstances properties *X* and *Y* make the same contribution to the causal powers of the things that have them, *X* and *Y* are the same property' (Shoemaker 1980: 256). It is crucial, then, to consider what relation mental properties bear to their realizers *vis a vis* the causal powers of the latter.

The view that mental properties, as second-order properties distinct from their first-order realizers, have causal powers of their own, has been challenged not only by Kim, but also by Ned Block (1990; 2003) and many others. Although it might seem that the discussion so far has just been concerned with a specific point of criticism about Kim's view on second-order properties and its implications for the generalization argument, I actually think that the above discussion suggests that realizer-functionalism is in a better position than role-functionalism for vindicating the causal status of mental properties, since realizer-functionalism straightforwardly identifies mental properties with their first-order realizers, so that the former 'inherit' the causal powers of the latter.

It is interesting to note that although Kim starts out with regarding mental/functional properties as second order properties, he seems to end up endorsing realizer-functionalism. Consider the following argument for the identification of M with its realizer P:

> So M is now the property of having a property with such-and-such causal potentials, and it turns out that property P is exactly the property that fits the causal specification. And this grounds the identification of M with P. M is the property of having some property that meets specification H, and P is the property that meets H. So M is the property of having property P. But in general the property of having property Q = property Q. It follows then that M is P.' (Kim 1998: 98-99).

To be sure, Kim himself recognizes that this identification is problematic. In the footnote appended at the very end of the quoted passage he says: 'The following question could have been occurred to astute readers … If M is a causal role and P its occupant, how could M and P be the same property? How could roles be identical with their occupants?' (ibid., note 11: 132).

However one deals with this subtle issue,[19] Kim's reductive strategy for identifying a mental property with its realizer is to employ what he has called the 'functionalization strategy'. The first step is to functionalize the property to be reduced by defining it in terms of its causal/nomic relations with other properties -- in a word, by specifying its causal role; the second step is to find the property in the reduction base that plays that role; the final step is to provide a theory that explains how the realizing property performs the specified role (Kim 1999; 2005).

Herein lies the problem: is the functionalization strategy compatible with the denial of *inter*level supervenience, given the need to search for a theory that explains how the realizer fills the causal role of the realized property? Or, to put the same question in other words: are the realizing and the realized properties necessarily at the same level, as required by Kim?

With respect to this question, the following responses are in order. First, we may rehearse the previous point concerning pain as a mental property of mine and the firing as a biochemical property of my C-fibers. The causal role for pain is for it to be caused by noxious stimuli and to cause winces and avoidance behaviours. This role is filled by my C-fibers firing, which is caused by impulses from sensory neurons in response to noxious stimuli and causes the activation of motor neurons resulting in bodily changes corresponding to winces and avoidance behaviours. It's true that all this happens *in me*, but the properties of my nervous system are *its* properties, not *mine*! Hence, the functionalization strategy presupposes interlevel supervenience.

The second response concerns the third 'theoretical' step of Kim's account of the functionalization strategy. As Kim is well aware, the process of finding out the occupant of a role (step 2) is intertwined with the process of theorizing (step 3); as he explicitly puts it elsewhere: 'Step 2 and 3 can be expected to be part of the same scientific research; ascertaining realizers of [a functionalized property] will almost certainly involve theories about causal/nomic interrelations among lower-level properties in the base domain'. (Kim 1999: 12) Consider genes and their realizer DNA (an example frequently used by Kim): Mendelian genetics tells us that there is an *x* whose role is to code and transmit information concerning phenotypic expressions; molecular biology – a lower-level theory – tells us that DNA is what plays that role. So it is within the framework of a theory that role properties can be reductively identified with the role-realizer properties. And, as the example illustrates, one needs to descend to a lower-level theory to find the realizer of the higher-level property.

Let me illustrate this with the help of another example. Consider a device like the one sketched below; let's call it a *v/s* (for 'valve/switch').

FIG 1 about here

This device is composed of two parts: a sustaining ring covered with an insulating (non-conductive) material, except for two diametrically opposed points (As) on it which

constitute the poles of an electric circuit connected to the ring, and a metallic disk that fits perfectly into the ring, and hinged to the ring (in points B) in such a way that when it is aligned with the ring it does not allow any substance to pass through. Now, suppose we place the v/s in a combustion engine as a butterfly or throttle valve: when the v/s is closed, the engine stops running ('is off'), when it is open the engine runs ('is on'). The v/s, however, can also be connected to an electric circuit, and then it works in the opposite way: when it is closed the circuit is on, when it is open it is off. This example shows that the first-order states of the v/s (the v/s being open or closed) are specifiable independently of their role in the v/s: the v/s is open when the disk is orthogonal to the sustaining ring, and close when it is aligned with it. But which functional role these first-order states occupy – i.e., which functional property they realize (if any) – can only be specified when it is specified whether the v/s is a component of an engine or of an electric circuit, and when the specification is in the context of a theory which explains how the entire system – the engine or the electric circuit – operates. Moreover, it is only in the context of such a theory that we recognize just these two states (open, closed) as the relevant states of the v/s. After all, there are indefinitely many physical states the v/s can be in – e.g. the hinged disk can be at indefinitely many different angles of rotation relative to the sustaining ring. These other states, while functionally irrelevant to the operation of some systems, would have been quite relevant to the operation of others, e.g., systems which the required an 'analogue' device like a butterfly valve with continuous states of partial opennes rather than a 'digital' on/off device admitting of no intermediate states.[20]

What the v/s case makes clear is that it is not possible to identify a first-order property of a system as the realizer of a 'role' property of the system without considering how it is able to fill that role in the context of an overall theory of the system. When the first-order properties are *microbased* properties (as in the case of pain, or in the case of Gillett's example about the cut diamond) the theory in question is, or includes, a lower-level theory about the *microproperties* of the system's proper parts. Unless the microbased properties (and their causal powers) were determined by, or supervened on, a mereological configuration of microproperties of the system's parts, it is difficult to see how those microbased properties could count as the realizers of a functional properties of the system's, or account for their causal powers. But to admit

this is, *pace* Kim, to admit interlevel supervenience, and this, in turn, is to legitimize the generalization argument. Surely, for Kim, an unwelcome result.

Acknowledgements

In writing this paper I greatly benefitted from advices and suggestions by Ausonio Marras, whom I warmly thank. I am also grateful to Jaegwon Kim for many helpful comments and suggestions on an earlier version of the paper. For further discussions on the ideas developed in this paper I thank Brian McLaughlin, Roberto Cordeschi and Mauro Dorato.

Simone Gozzano
Dipartimento di Storia e metodologie comparate, Università di L'Aquila, via Roma 33, 67100 L'Aquila, ITALY
simone.gozzano@cc.univaq.it

References

Achinstein, P. (1974), 'The Identity of Properties', *American Philosophical Quarterly* 2: 257-276.

Baker, L.R. (1993), 'Metaphysics and Mental Causation' in J. Heil and A. Mele (Eds) *Mental Causation*, Clarendon Press: Oxford.

Block, N. (1980), 'Introduction: What Is Functionalism?' in N. Block, (Ed) (1980) *Readings in the Philosophy of Psychology*, 2 voll., Harvard University Press: Cambridge MA, vol. 1.

Block, N. (1990), 'Can the Mind Change the World?' in G. Boolos (Ed) *Meaning and Method*, Cambridge University Press: Cambridge.

Block, N. (1997), 'Anti-Reductionism Slaps Back', *Philosophical Perspectives* XI: 107-32.

Block, N. (2003), 'Do Causal Powers Drain Away?', *Philosophy and Phenomenological Research* 67: 110-127.

Bontly, T. (2002), 'The Supervenience Argument Generalizes', *Philosophical Studies* 19: 75-96.

Burge, T. (1993), 'Mind-Body Causation and Explanatory Practice' in J. Heil and A. Mele (Eds) *Mental Causation*, Clarendon Press: Oxford.

Crisp, T. and Warfield, T. (2001), 'Kim's Master Argument', *Noûs* 35: 304-316.

Dennett, D.C. (1978), *Brainstorms*, MIT Press, Cambridge MA.

Dennett, D.C. (1987), *The Intentional Stance* MIT Press, Cambridge MA.

Gillett, C. (2002), 'The Dimensions of Realization', *Analysis* 62: 316-23.

Heil, J. (2003), 'Levels of Reality', *Ratio* 16: 205-221.

Jackson, F., Pargetter, R. e Prior, E. (1982), 'Functionalism and Type-type Identity Theories', *Philosophical Studies* 42: 209-25.

Kim, J. (1997), 'Does the Problem of Mental Causation Generalize?', *Proceedings of the Aristotelian Society* 97: 281-97.

Kim, J. (1998), *Mind in a Physical World*, MIT Press: Cambridge MA.

Kim, J. (1999), 'Making Sense of Emergence', *Philosophical Studies* 95: 3-36.

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

Kim, J. (2003), 'Blocking Causal Drainage and Other Maintenance Chores with Mental Causation' *Philosophy and Phenomenological Research* 67: 151-176.

Kim, J. (2005), *Physicalism, or Something Near Enough*, Princeton University Press: Princeton.

Lewis, D. (1966), 'An Argument for the Identity Theory' *Journal of Philosophy* 63: 17-25

Lewis, D. (1972), 'Psychophysical and Theoretical Identifications', *Australasian Journal of Philosophy* 50: 249-58.

Lewis, D. (1980), 'Mad Pain and Martian Pain' in N. Block, (Ed) (1980) *Readings in the Philosophy of Psychology*, 2 voll., Harvard University Press: Cambridge MA, vol. 1.

Loar, B. (1981), *Mind and Meaning*, Cambridge University Press: Cambridge.

Lycan, G. (1987), *Consciousness*, MIT Press, Cambridge MA.

Marras, A. (2000), 'Critical Notice of *Mind in a Physical World*', *Canadian Journal of Philosophy* 30: 137-60.

Millikan, R.G. (1984), *Language, Thought and other Biological Categories*, MIT Press: Cambridge MA.

Noordhof, P. (1997), 'Making the Change: the Functionalist's Way', *British Journal for Philosophy of Science* 48: 233-250.

Shoemaker, S. (1980), 'Causality and Properties', reprinted in J. Kim and E. Sosa (Eds) *Metaphysics. An Anthology*. Blackwell: Oxford.

Shoemaker, S. (1981), ‘Some Varieties of Functionalism’, reprinted in S. Shoemaker *Identity, Cause and Mind. Expanded edition*, Oxford University Press: Oxford.

Shoemaker, S. (2005), ‘Realization and Mental Causation’, reprinted in S. Shoemaker *Identity, Cause and Mind. Expanded edition*, Oxford University Press: Oxford.

Sober, E. (1999), ‘The Multiple Realizability Argument Against Reductionism’, *Philosophy of Science* 66: 542-564.

Van Gulick, R. (1993), ‘Who’s in Charge Here? And Who’s Doing All the Work?’ in J. Heil and A. Mele (Eds) *Mental Causation*. Clarendon Press: Oxford.

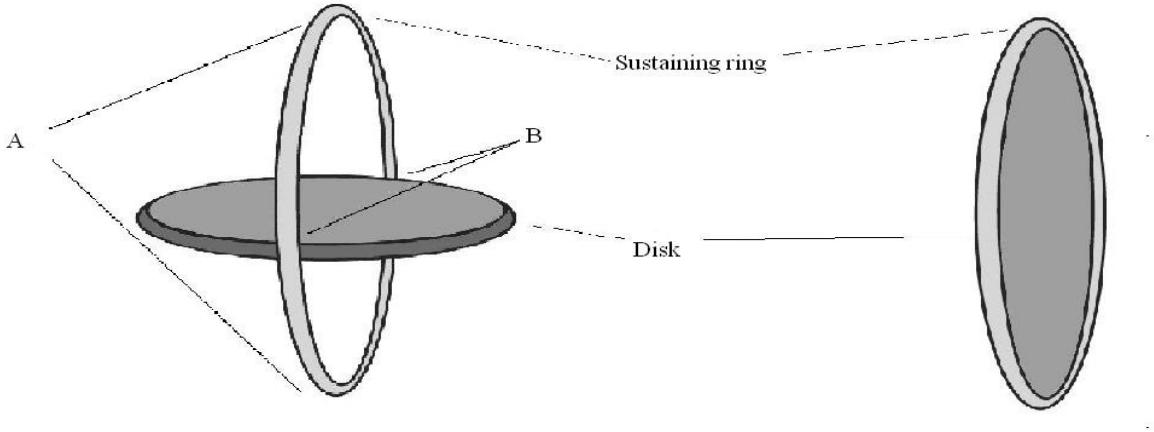

Fig. 1

Notes

[1] These two varieties of functionalism correspond to what Ned Block calls 'functional specification' and 'functional state' functionalism, respectively (Block 1980: 179). Brian McLaughlin uses the expression 'filler functionalism' to refer to what it is more commonly called 'realizer functionalism'; I will adopt the latter expression. Sometimes even 'role functionalism' is, unfortunately, used for the same purpose.

[2] In favor of overdetermination see Crisp and Warfield (2001). But see Kim reply in (2005: 46-52).

[3] At least of the 'role' variety (cf. Block 1980). More on the 'role-' vs. 'realizer-' functionalism distinction below.

[4] Actually Kim characterizes the further condition as the condition of being either red or blue or green. But this condition merely provides a list of the condition's satisfiers rather than a general specification of the satisfaction condition itself.

[5] If, as Kim holds (1998: 119), a condition for being a distinct property is having distinct causal powers, then second-order properties must either be identified with their first-order realizers on each occasion, or, if such an identification is disallowed, they must be regarded as not real properties at all, and talk about second-order properties would best be replaced by talk about second-order *concepts* (cf. Kim 1998: ch. 4). However this may be, the point is that second-order properties have no causal powers over and above those of their realizers on each occasion. It is thus unsurprising that for Kim the same should be true of mental properties.

[6] Block (1997; 2003), Bontly (2002), Marras (2000). Earlier claims to the same effect can also be found in Burge (1993), Baker (1993), Lycan (1987), Noordhof (1997), van Gulick (1992).

[7] See also Kim (1998: 112) for his own reconstruction of the debate. Block (1997) also considers the possibility that there is no final level. I leave this aside.

[8] For Kim, micro-based properties are properties that things have in virtue of their being completely decomposable into non overlapping parts possessing certain microproperties and standing in an appropriate relation to one another (cf. Kim 1998: 84). The important point is that micro-based properties, unlike microproperties, are possessed by wholes and not by their parts, and so such properties are at the same level as the properties they realize. Being a water molecule is Kim's example of a micro-based property; its being composed of two hydrogen atoms and one oxygen atom in the appropriate bonding relation is a property of the whole molecule, not a property of its parts (the respective atoms).

[9] This, of course, is a direct consequence of the definition of supervenience that is operative in the supervenience argument – a definition that requires supervenient properties and their base properties to be properties *of the same object*. Bontly (2002) refers to this requirement as the 'coinstantiation' requirement.

[10] See also Kim (1997: 291).

[11] It may be relevant to note that, in general, the question whether the properties of a whole system are or are not shared by its parts is independent of the question whether they supervene on the properties of the system's parts. A 'summative' property like the mass of this table, though determined by the mass if its parts, is not shared by the table's parts (e.g. by its top or by its pedestal), whereas a non-summative, 'distributive' property like a certain rate of acceleration is shared by a system and its parts. Nonetheless, both the mass and the acceleration of a system are determined by, and supervene on, the mass and acceleration of its parts: changes with respect to the former necessitate changes with latter; fixing the latter, fixes the former. Mereological supervenience, then cuts across the problem of levels: supervenience may hold between properties applying to objects at different levels as well as at the same level.

[12] Assuming that there is such a level; cf. Block (2003) for discussion.

[13] And they remain so even in so-called cases of multiple realization, where a mental property is relativized to a species (or structure type), and each relativized mental property gets identified with the occupant of its role for each distinct species or structure type (cf. Lewis 1980).

[14] As Loar puts it: 'The states over which a functional theory quantifies are not themselves functional states … They are *first–order* properties, properties of individuals … A *functional state* of an individual is a second-order state or property – being in a first-order state with a certain functional role' (Loar 1981: 45)

[15] Unsurprisingly, role-functionalism is the version of functionalism favored by the nonreductive physicalists.

[16] The level in the hierarchy at which we find the micro-base for a mental (or other second-order) property might vary with our explanatory interests and with the diversity of the property's possible realizers in the same or in different species.) Whether mental properties find their micro-base at the neural, biochemical, or physical level, and whether they may have multiple bases in different organisms is an issue that shall not concern us here.

[17] Such as homuncular functionalism, (Lycan 1987; Dennett 1978, 1987) or teleological functionalism (Millikan 1984).

[18] See Achinstein (1974) for an earlier formulation.

[19] Kim's favorite way of dealing with this issue seems to be to turn role properties into role *concepts* (Kim 1998: 103-104).

[20] It is interesting to consider the v/s device in the light of Shoemaker's distinction between the *core* and the *total* realizer of a given functional property (Shoemaker 1981; 2005). The core realizer is the physical state of the embedded v/s; the total realizer is the physical state of the whole system in which the v/s is embedded. In order to determine what functional property the v/s is realizing we need to consider the state of the whole system. Analogously for the 'pain' case: the C-fibers firings in a properly 'wired' brain are the core realizer of pain; the whole brain (or nervous system) in which the firing C-fibers are embedded is the total realizer. Perhaps David Lewis had something similar in mind when he stated that 'the concept of pain, or indeed of any other experience or mental state, is the concept of a state that occupies a certain causal role, a state with certain typical causes and effects. … It is the concept of *a member of a system of states* that together more or less realize the pattern of causal generalizations set forth in commonsense

psychology. (*That system may be characterized as a whole and its members characterized afterward by reference to it* .)' (Lewis 1980: 218; my emphasis)