

FUNCTIONALISM AND NONREDUCTIVE PHYSICALISM†

David PINEDA*

* Departament de Filologia i Filosofia, Facultat de Lletres, Universitat de Girona, Pl. Ferrater Mora 1, 17071 Girona, Spain. E-mail: david.pineda@udg.es

BIBLID [0495-4548 (2001) 16: 40; p. 43-63]

ABSTRACT: Most philosophers of mind nowadays espouse two metaphysical views: Nonreductive Physicalism and the causal efficacy of the mental. Nevertheless, this position is threatened by a number of serious difficulties. In this paper, I propose a metaphysical account of functional properties and show how this proposal is able to overcome some of these difficulties, in particular, some recent arguments against the causal efficacy of multiply realized properties. However, in the second part of the paper an objection against this proposal is raised and, after a detailed discussion of it, the conclusion reached is that the prospects for a functionalist nonreductive metaphysics of the mind which affords causal powers to the mental seem certainly dim.

Keywords: mental causation, Nonreductive Physicalism, Functional properties, multiple realization.

CONTENTS

1. The proposal
 2. The Problem
- Bibliography

Most philosophers of the mind nowadays espouse two metaphysical views: Nonreductive Physicalism and the causal efficacy of the mental. Throughout this discussion I will refer to the conjunction of both claims as the Causal Autonomy of the Mental. Nevertheless, this position is threatened by a number of difficulties which are far more serious than one would imagine given the broad consensus that has been reached in recent decades. This paper purports to offer a careful examination of some of these difficulties and to show the considerable efforts that have to be made in order to try to overcome them. The difficulties examined concern only those metaphysical problems common to all special science properties but not those that are specific of mental properties. So, in proposing a functionalist version of Nonreductive Physicalism in what follows, I will not attempt to answer well-known objections such as the absent qualia argument and the like. This should not be interpreted as a limitation in the scope of this work. On the contrary, in dealing with more general

objections we will try to evaluate a position which entails (under common assumptions) the Causal Autonomy of the Mental, namely: Nonreductive Physicalism plus the causal efficacy of special science properties.

The aim of this work is twofold, corresponding to its positive objectives and to its negative aims. The positive undertaking is the proposal of a metaphysics of functional properties that is free of some of the objections to the Causal Autonomy of the Mental; the negative undertaking consists in the consideration of further problems which will cast doubt on our previous efforts to render plausible the Causal Autonomy of the Mental. The conclusion drawn is that this view, its overwhelming acceptance notwithstanding, is more problematic than it might otherwise seem at first sight.

1. The proposal

The central tenet of the Causal Autonomy of the Mental could be summarised by the following proposition:

- (1) Mental properties (and, in general, special science properties like biological or geological properties) appear in causal laws and are multiply realized by physical properties.

Many assumptions are implicit in (1) and these will be spelled out in a moment. However, let me make two remarks before I try to tackle them. The first remark is that in my opinion nothing essential hinges on these assumptions. They are important because they will help us to shape the discussion that will follow, but I do not believe that any of the problems discussed depend on them. My second remark is that all of them are assumptions endorsed by the vast majority of the defenders of the Causal Autonomy of the Mental, so that no extra assumptions not taken into account in the current debate will be introduced.

Let's turn our attention to the assumptions. One of them states that the empirical world is governed by laws; another states that laws consist partly of properties, and in particular that causal laws are constituted by a certain relation (which we do not need to analyse here for present purposes) between antecedent or cause-properties on the one hand and consequent or effect-properties on the other. Given these views on laws, the claim that an instantiation of a mental property *M* is causally efficacious with respect to an instantiation of another property *P* implies that there exists a causal law such that *M* is one of its antecedent properties and *P* is one (or the only one) of its consequent properties. Moreover, to say that a property *M* is multiply realized by a class of properties *C* implies that each instantiation of *M* occurs in virtue of the instan-

tiation of a property belonging to C and that different instantiations of M can occur in virtue of instantiations of different properties in C.

Assuming that the instantiation of a law consists essentially in the instantiation of the properties constituting it, the picture that emerges from the Causal Autonomy of the Mental together with these assumptions is the following. Any instantiation of a special science law, i.e., any instantiation of a biological, geological or psychological law, requires -in the sense that occurs in virtue of- the instantiation of a physical law, constituted partly by the physical properties which are the realizers in that occasion of the special science properties constituting the special science law. Given that the latter properties are multiply realized by physical properties, we cannot identify the special science law with any physical law (as Reductive Physicalism would recommend us to do), since different instantiations of the same special science law would take place in virtue of instantiations of different physical laws. So, here we have a sense in which physical laws are basic, since any other causal law depends on them. But they are not the only ones, since there also exist causal laws -biological, geological, psychological laws-, which are not physical. In sum, here we have a position which is compatible with orthodox physicalism and affords full metaphysical autonomy to the special sciences, which is the crucial tenet of Nonreductive Physicalism.

Before starting to discuss the problems of this approach, let me say at the very outset that I take the Causal Autonomy of the Mental to be the most commonsensical view about the mind. First, because of the empirical evidence of the dependency of the mental on the physical that is being built up. Second, because denying the causal efficacy of the mental goes against some of our most firm convictions (Would Ethics or Law make sense were the mental causally idle?). Finally, because the denial of the multiple realizability claim seems preposterous: are we seriously willing to assert that all organisms capable of feeling pain instantiate the same physical property when they experience pain? (Putnam 1967 and 1973). However, even more commonsensical views have been refuted by means of argument, e.g., the view that the sun moves around the Earth. Hence, to acknowledge the commonsensical character of the Causal Autonomy of the Mental does not by itself guarantee its truth and neither does it entitle us to ignore its difficulties.

The first difficulty I want to explore is that of making clear the key relation that holds between properties according to which one property realizes another in such a way that different properties may realize one and the same property. Without a clear understanding of that relation we will not have a clear understanding of the possibility of the mental being "multiply realized"

by the physical either, nor of the claim of the Causal Autonomy of the Mental. We need to dispel the sense in which an instantiation of a property *occurs in virtue of* the instantiation of another property, implicit in the realization idiom. Or, what amounts to the same thing, we need to make it clear in exactly what sense a special science law depends on physical laws though at the same time remaining autonomous of them.

However, this task is far from easy, as we shall presently see. Some nonreductive physicalists worried by this problem proposed some time ago to elucidate the realization relation among properties as a kind of supervenience relation. The favourite was strong nomic supervenience (Kim 1984), though it was not the only one (others advocated global nomic supervenience). But there seems to be powerful objections against this proposal. First, a supervenience relation between properties (either strong or global) does not look like a dependence relation between them, for at least two reasons: a) supervenience relations, as they are usually defined, are never asymmetrical, whereas a dependence relation should be asymmetrical -if a property P depends on a property Q then it cannot be that the property Q also depends on the property P. At least not in the sense of dependence required by the Causal Autonomy of the Mental. Nonreductive physicalists want mental properties to be realized by physical properties in a sense in which it cannot be the case also that the latter are realized by the former; b) a relation of supervenience is simply a relation of covariance among properties. To say that mental properties covary in certain ways with certain physical properties is very different from saying that they become instantiated in virtue of the instantiation of those physical properties. One can, in effect, endorse covariance, supervenience, and reject dependence of the right physicalist sort. Furthermore, the claim of the supervenience of the mental on the physical is too weak for the nonreductivist since it is compatible with certain forms of Dualism. For example, one could hold that mental properties correlate with physical properties in such a way that a relation of supervenience holds among them, but the former are not dependent on the latter (such a position was defended by prominent dualists including Malebranche and Leibniz).

Some physicalists, in light of these or similar reasons, consider that the realization relation entails but is not entailed by a supervenience relation and therefore cannot be analysed in terms of it (Kim 1998).

The question that suggests itself is: what other relation could do the job? My first proposal in this paper is that in order to answer this question one must bear in mind what I take to be a central requisite of the notion of realization of a property by another. The requisite is that there must be an explanatory

link between realizers and realized properties in such a way that each instantiation of the realized property can be explained in terms of the instantiation of its realizer on that occasion. The reason for that requisite is that in the absence of this explanatory link there is no way of distinguishing realization from mere covariance. Given that we do indeed distinguish these two sorts of relation among properties, therefore I conclude that the idea of an explanatory link is part of the concept of realization involved in the claim of the Causal Autonomy of the Mental. So in saying that a particular instantiation of a property P realizes a particular instantiation of a property Q we imply that there is an explanation of the latter instantiation in terms of the former.

Some restrictions as to the kind of explanation required here are in order. Quite obviously one of them is that mention of the instantiation of the realizer property must be crucial in the explanation. Another important restriction is that in the explanation there cannot be any appeal to properties or laws belonging to the same ontological level of the realized property. This second restriction is particularly clear in the case of the multiple realizability of mental properties: if we could not explain a particular instantiation of a mental property M without appealing to some mental properties or mental laws, then there would be empirical mental facts not wholly accountable in terms of physical facts, and quite independent of them, and that would be contrary to the claim of Nonreductive Physicalism.

These restrictions are hard to meet. And it is here, in my opinion, where Functionalism proves to be of great help for the defender of the Causal Autonomy of the Mental. This is so because functional properties, at least from a particular understanding of them, hold with their realizers an explanatory relation that meets the restrictions of realization. My proposal is then to explore a functionalist version of Nonreductive Physicalism and to see whether it can guarantee the causal efficacy of the mental in which case we would have a solid defence of the Causal Autonomy of the Mental.

Functional properties are often characterised by the following proposition:

- (2) Something X instantiates a functional property F if and only if X instantiates any property having the functional role associated with F. (Schiffer 1987).

'functional role' in (2) is used to mean a property of properties, more precisely, the property instantiated by some properties when and only when they enable those entities instantiating them to cause and be caused in the ways specified in the functional role. Since we will be dealing with very complex problems I will focus only on functional properties with a simple functional

role associated with them, properties such as solubility in water and dormitivity. Thus, according to (2), we will say that something X instantiates solubility in water if and only if X instantiates some property which causes X to dissolve when submerged in water. Likewise, we will say that something X instantiates dormitivity if and only if X instantiates any property causing sleep.

A realizer of a functional property F, according to this characterization, is a property having the functional role associated with the functional property. Therefore, a realizer of solubility in water is any property causing any entity instantiating it to dissolve when submerged in water (and normal conditions obtain); and a realizer of dormitivity is any property causing sleep (in normal conditions). It is clear that some functional properties will have more than one realizer, that is, that in some cases more than one property is going to have the right functional role. This is the case, in our example, of dormitivity, since more than one property has the property of causing sleep.

Let us see now whether the relation between a functional property and its realizers meets the requirements of the realization relation which is of interest to the defender of the Causal Autonomy of the Mental. The characterization offered in (2) of functional properties makes perspicuously clear their explanatory link with their realizers, since (2) entails that the instantiation of a functional property requires the instantiation of one of its realizers. Furthermore, we have an explanation of each instantiation of a functional property F in terms of instantiations of its realizers. In fact, we explain a particular instantiation of F by the instantiation of some property having the functional role associated with F. For example, we can explain a particular instantiation of solubility in water by something X through the instantiation by X of a property P which causes the dissolution of X when submerged in water. It goes without saying that mention of the realizer is crucial in this explanation and, furthermore, that there is no appeal to properties or laws not belonging to the level of the realizer itself. Hence, as far as the requisites of the realization relation are concerned, we could in principle elucidate it as the sort of relation held by a functional property and the properties having the functional role associated with it.

So we could defend the Causal Autonomy of the Mental as follows. We could say that mental properties, and other special science properties like biological or geological properties, are functional properties in the sense that they are individuated by a functional role, and that they are multiply realized by physical properties in the sense that there is, in some cases, more than one physical property with the same functional role. This proposal is however still insufficient for the purpose of defending the Causal Autonomy of the Mental.

What we have just done is simply to defend the claim of Nonreductive Physicalism by analysing the realization relation mentioned in (1) as the relation of functional realization. But we still have to argue for the causal efficacy of the mental. What I want to do next is to defend this claim from some powerful objections which are, in fact, targeted at the more general thesis of the causal efficacy of any special science property multiply realized by physical properties. In order to do that we will have to endow our metaphysical proposal with some more complexity.

Indeed, to answer some of the objections we need to say something more about the nature of functional properties. Notice that (2) gives us only the conditions under which it is correct to attribute to an entity a functional property -namely, when that entity instantiates a property having the right functional role-, but it tells us nothing about what sort of thing a functional property is, or to put it in other words, it does not tell us exactly what a functional property consists in.

There are two positions with regard to functional properties which can be adopted as two initial answers to this last question. These two positions are summarized by the following two propositions:

- (3) for any functional property F and for any instantiation *i* of F, *i* is the instantiation of its realizer in that occasion. (Lewis 1994; Kim 1997).
- (4) for any functional property F and for any instantiation *i* of F, *i* requires the instantiation of some realizer of F, but is not the same entity as that instantiation.

To distinguish between these two positions, I will refer to the one characterized by (3) as the first order interpretation of functional properties, and to the one characterized by (4) as the second order interpretation of functional properties. I call the latter the "second order interpretation" because it characterizes any instantiation of a functional property by quantifying over the instantiation of another property.

Notice that the first interpretation tells us more about the nature of functional properties than the second, since it tells us what their instantiations consist in. The second one, in contrast, tells us what their instantiations depend on, but not what they consist in. However, in my opinion there is a knock-down argument against the causal efficacy of functional properties understood under the first order interpretation, whence the defender of the Causal Autonomy of the Mental needs to turn to the less precise second order interpretation of functional properties.

Here, briefly, is the argument, inspired by one consideration in Kim (1992):

(5)

- (i) All the instantiations of a causally efficacious property must share the same causal powers.
- (ii) Not all the instantiations of a multiply realized functional property share the same causal powers.

 (c) Multiply realized functional properties are not causally efficacious properties.

This argument makes use of the notion of the causal powers of an instantiation of a property. Although the use of this notion is ubiquitous in the literature its meaning is seldom elucidated. To discuss this argument thoroughly, however, we will have to tackle this task. Intuitively speaking, the causal powers of a given instantiation of a property P consist in this feature of such instantiation in virtue of which the entity having it is capable (if appropriate conditions hold) of undergoing certain causal transactions. Now, assuming that causal relations are backed by causal laws (an assumption commonly shared by the defenders of the Causal Autonomy of the Mental, which among other things endorse the existence of mental causal laws), then an instantiation of a property P has the causal powers that it has in virtue of being an instantiation of a causally efficacious property, that is, of a property figuring in causal laws. This being so, it follows then that the premise (i) of the argument (5) is unquestionably true. The reason for this is as follows. Take two different instantiations, *i* and *j*, of a property P. Since causal transactions are backed by causal laws and both *i* and *j* are instantiations of the same property P, it follows that the causal transactions enabled by *i* and the causal transactions enabled by *j* are both backed by causal laws in which P figures and hence that the set of causal transactions enabled by *i* is the same as the set of causal transactions enabled by *j*. Therefore, *i* and *j* share the same causal powers. Given this result I will henceforth, to avoid verbosity, speak of the causal powers of a property instead of the causal powers of a given instantiation of a property.

Now, a very simple reasoning shows that premise (ii) of argument (5) is also true, under the first order interpretation of functional properties. Suppose that a functional property F has at least two different realizers P and Q. In principle, P and Q may or may not have different causal powers (there is no need here to assume the view that a property is individuated by its causal powers, in which case P and Q being by supposition different should also have dif-

ferent causal powers). However, for the defender of the Causal Autonomy of the Mental who wants to argue that mental laws do exist that are autonomous with respect to physical laws, and who also wants to contradict the reductivist claim that mental laws simply are, or are correlated to, physical laws, it is crucial that mental properties have physical realizers with different causal powers. Suppose then, further, that P and Q have in effect different causal powers. Now take two different instantiations *i* and *j* of F and suppose finally that in the case of *i* F is being realized by an instantiation of P, and in the case of *j* F is being realized by an instantiation of Q. The causal powers of the instantiation of P being different from the causal powers of the instantiation of Q, and *i* being the instantiation of P and *j* being the instantiation of Q (under the first order interpretation of functional properties), then by Leibniz's law, the causal powers of *i* are different from the causal powers of *j*.

Since argument (5) is obviously valid and its two premises have been shown to be true, then conclusion (c) holds if we understand functional properties in terms of the first order interpretation. Does conclusion (c) also hold if we wished to adopt the second order interpretation of functional properties? Were it to do so, we would then have to abandon our functionalist proposal to defend the Causal Autonomy of the Mental. But I think it does not, although to demonstrate this we would have to complement once more our account of functional properties.

J. Kim has proposed (in Kim 1992; 1993) a metaphysical principle which, were it true, would render premise (ii) of argument (5) true regardless of how functional properties are understood. The principle, called the Causal Inheritance Principle, reads as follows:

- (7) If a property F gets instantiated in virtue of being realized by another property P, then the causal powers of that instantiation of F are identical to the causal powers of P.

Now, if Kim's principle is true, then even if, according to our second order interpretation of functional properties, an instantiation of F is not identical to the instantiation of its realizer on that occasion, both instantiations share all their causal powers. Therefore, two instantiations of F which occur by being realized by two realizers of F having different causal powers will also have different causal powers, and conclusion (c) will follow (assuming (i), of course).

Notice that Kim's principle is quite general, because it concerns any multiply realized property. It holds true no matter how the realization relation involved in it is elucidated. The generality of the principle is in my opinion the

cause of its debility, since in the absence of a clear account of the relation in which two properties stand when one is realized by the other (an account which Kim does not bother to offer), there seems to be no way of justifying (7). The principle could be true or not depending on how the realization relation is understood in general and on the nature of the realization relation proposed. I will show that the principle is false in some cases, and particularly false if the relation of realization proposed is the relation of functional realization and functional properties are understood under the second order interpretation.

What then are Kim's reasons for such a general principle? As far as I can see, his reasons are contained in the following paragraph:

I challenge these nonreductivists who would reject this principle to state an alternative principle on just how the causal powers of a realized property are connected with those of its realization base; or explain, if no such connections are envisioned, the significance of the talk of realization (Kim 1993, p. 355).

In my understanding of this passage Kim is making two points. The first point is that to say that a property P is realized by a property Q implies that there is some "connection" between the causal powers of P and the causal powers of Q. The second point is that the only connection possible is that mentioned in principle (7). Now, I agree with Kim on the first point, but disagree on the second. Note, to begin with, that if Kim were right it would be impossible to defend the Causal Autonomy of the Mental provided that one also wants to defend the claim that causal relations are backed by causal laws. The causal autonomist of the mental requires two claims to be true: mental properties are multiply realized by physical properties (with different causal powers) and mental properties are causally efficacious (with respect to other properties), and these two claims are incompatible if Kim's second point is true. Hence, Kim is arguing not merely for the falsity of the Causal Autonomy of the Mental but for the *unintelligibility* of such a position. The best way to reply to Kim's objection, therefore, is to offer some intelligible version of the Causal Autonomy of the Mental. This is what I will try to do next.

I will begin by arguing in favour of Kim's first point. We should recall that Nonreductive Physicalism obviously purports to be a form of Physicalism. The core idea in Physicalism is that every empirical entity and process depends on physical entities and processes, including causal processes. Now, if any causal process depends on physical (causal) processes, then it cannot be the case that the causal powers of, say, mental properties, are independent of, or unconnected with, the causal powers of physical properties. Therefore, if the claim that mental properties are realized by physical properties is to count as a physicalist claim (as the causal autonomist of the mental would have it), then

it must be true that the causal powers of realized properties are dependent on the causal powers of their realizers. Otherwise the claim that mental properties are multiply realized by physical properties would not be a physicalist claim.

So then, what the defender of the Causal Autonomy of the Mental must respect is the following principle, and not the principle proposed by Kim:

- (8) For any realized property P, the causal powers of any instantiation *i* of P depend on the causal powers of the realizer of P on the occasion of *i*.

This principle is obviously satisfied by (7), but there are other proposals of the relation between the causal powers of realized properties and their realizers as regards (8) other than identity. I want to offer such a proposal.

The leading idea is expressed in the next proposition:

- (9) The causal powers of a functional property are individuated by the properties mentioned in the functional role associated with it.

Notice, first of all, that given our key assumption about the nomologicity of causation and the role of properties in causal laws, it is fairly natural to individuate causal powers in terms of properties. This means that the notion of property occurs prior to the notion of causal power, but for the sake of exposition, I will assume this prioricity here without argument. I think that nothing essential hinges on this question and, in any event, there are surely other ways of specifying my proposal without relying on this assumption.

Going back to our examples of functional properties, it turns out, according to (9), that the causal powers of solubility in water are individuated partly in terms of the property of dissolution in water, since that property is mentioned in the specification of the functional role associated with solubility in water. Likewise, the property of sleep will be one of the properties in terms of which the causal powers of dormitivity will be individuated.

This proposal about the causal powers of functional properties is not "ad hoc". The main rationale for it is the characterization (2) of properties of this kind. According to this characterization, what is specific about a functional property is precisely that it has a functional role associated with it, but this functional role is nothing more than a pattern of causal relations. So, if we identify a property by a pattern of causal relations and we want to defend the causal efficacy of such a property, it is very reasonable to hold it causally efficacious with respect to the properties involved in that pattern. That is precisely what proposal (9) seeks to do.

Moreover, notice that all of the realizers of a given functional property must be causally connected with the properties mentioned in the functional role associated with the functional property in order to count as such realizers.

Our simple examples show this point very clearly. For example, any realizer of solubility in water must cause dissolution in water of that which instantiates it (under normal conditions). The modality here is conceptual, for the existence of this causal connection is what qualifies it as a realizer of solubility in water. Likewise, any realizer of dormitivity must cause sleep, for similar reasons. Hence, our proposal about the causal powers of a functional property has the following interesting consequence:

- (10) The causal powers of a functional property F are a subset of the causal powers common to all realizers of F.

This result follows directly from our characterization (9) of the causal powers of a functional property together with the fact that any realizer of a functional property F has the functional role associated with F. All realizers of F must at the very least have in common the causal powers in virtue of which they qualify as realizers of F, and according to our characterization, these causal powers must also correspond to F.

The relation exhibited in (10) between the causal powers of a functional property and the causal powers of its realizers also offers us a rejoinder to Kim's argument concerning the causal efficacy of multiply realized properties. As (10) makes clear, our proposal respects the general physicalist principle that the causal powers of a functional property are dependent on the causal powers of its realizers. However, any two instantiations of a functional property will share the same causal powers according to our proposal. This is the case since any two instantiations of the same functional property are associated with the same functional role and causal powers are individuated by relation to the functional role. Finally, the causal powers of a functional property F need not be the same as the causal powers of its realizers, certainly not if F, as the case might be, has two or more realizers with different causal powers. This will ensure the metaphysical autonomy of functional properties and show that principle (7) is false. So then, any two instantiations of a functional property will share the same causal powers, and be different from the causal powers of its realizers. So far, it seems that we have met Kim's challenge, and have at our disposal a nonreductive metaphysics capable of supporting the claim of the Causal Autonomy of the Mental.

2. *The Problem*

Our proposal should be refined in a number of ways to account for the causal powers of functional properties with more complex functional roles associated with them than those associated with dormitivity and solubility in water.

Thus, some functional roles will involve other functional properties. For example, if we find plausible the view according to which the property of believing that P is a functional property, then we will probably also find plausible that the functional role associated with it includes that believing that P together with believing that Q causes in the object which instantiates those properties the further property of believing that P and Q. Hence, to count as a realizer of believing that P any such realizer must cause, together with a realizer of believing that Q, a realizer of believing that P and Q, under normal conditions. It seems natural to think that if we want believing that P to be a causally efficacious property it must be efficacious with respect to believing that P and Q, and so the capacity of bringing about the latter property must be among its causal powers. To cope with this case we should, as I said, make our proposal more sophisticated. Still, our characterization (9) would hold, but it is dubious whether the consequence (10) would follow as well. There could be independent reasons not to regard a realizer of believing that P as causally efficacious with respect to believing that P and Q, and in that case the causal powers of believing that P could not be a subset of the causal powers common to its realizers, being the bringing about of believing that P and Q among the former but not among the latter. In any event, whether we want to regard a realizer of believing that P as causally efficacious with respect to believing that P and Q or not, the causal powers of believing that P would still depend, in the right physicalist way, on the causal powers of their realizers, since an instantiation of believing that P and Q would in any case require an instantiation of one of its realizers which in turn would be caused partly by the realizer of believing that P. The validity of this dependency claim, together with the falsity of the Causal Inheritance Principle in this case too, is all that we need to reply to Kim's sceptical argument.

However, Kim has another sceptical argument against the causal efficacy of multiply realized properties which would apply to our proposal. The argument is complex and deserves a careful examination, but the gist of my position on it can be summarized in a few paragraphs. Notice, to begin with, that our characterization of the causal powers of functional properties commits us to massive cases of "downward causation", as Kim would put it (Kim 1993). Indeed, it follows from our characterization that the instantiation of certain properties (namely, those mentioned in the functional role associated with the functional property) will be caused by both the instantiation of the functional property and the instantiation of its realizer, whereby the name of "downward causation", since we claim a causal relation of a property F with another property causally related to one of the realizers of F, belonging therefore to a more

fundamental ontological level. Kim holds that such a claim is absurd in the light of the Principle of the Causal Closure of the Physical Domain, according to which the instantiation of physical properties is causally sufficient for the instantiation of a physical property. He argues that, given that the physical realizers are causally sufficient to bring about the physical effect, what entitles us to hold the functional properties causally responsible for that physical effect also? (Kim 1993; 1998).

This question, which is usually referred to in the literature as "the causal exclusion problem", has a short and direct answer: common sense. Let me expand on this answer a little. It has been repeatedly observed that Kim's argument, if valid, would render causally efficacious only fundamental physical properties. So any properties not reducible to them would be deprived of causal powers. Now, common sense attributes causal powers to certain properties, like mental, geological or biological properties, which it also tends to regard as not physically reducible at all.

Is this answer good enough for the defender of the Causal Autonomy of the Mental? The answer is based on the implicit premise that any argument having a consequence contrary to common sense is incorrect. Now, as a general principle this premise is false in scientific disciplines. Otherwise, it would rule out heliocentrism, to give a simple example. It should then be argued that it is true in the case of philosophical disciplines, in particular metaphysics, but that surely requires a long and complex argumentation.

Be it as it may, it should be mentioned that Kim's considerations, even if we decide to accept them fully, only show that special science causation is redundant, not fundamental causation, and they do not demonstrate, on their own, that it is no causation at all. In fact, this "exclusionary" conclusion should come as no surprise for those of us who have characterized the causal powers of functional properties as a subset of the causal powers common to all of their realizers. Redundancy was apparent here in the characterization itself. And indeed, it should not surprise us that physicalist considerations, such as those invoked by Kim, lead us to the conclusion that only physical causation is fundamental and that special science causation is at best redundant. This is simply what we should expect to happen as a result of embracing the physicalist principle (8) above. If the causal powers of special science properties should depend on the causal powers of their realizers, then it is perfectly coherent to hold that only physical properties have fundamental, non redundant causal powers.

However, our metaphysical proposal in support of the claim of the Causal Autonomy of the Mental is still incomplete in one important respect: the na-

ture of the functional properties. Earlier, we rejected the first order interpretation of functional properties and adopted the second order interpretation. As we said there, this interpretation is very imprecise about the nature of an instantiation of a given functional property. It only tells us that it requires the instantiation of a realizer of that functional property, but nothing more than that, and in particular nothing about how we are to construe our functional properties.

To show the importance of this limitation, and the difficulty of overcoming it, I will discuss a further argument against the proposal defended in the last section, an argument which has been cropping up in the literature in recent years, and which I call "the modal argument". The argument, inspired by that offered in Fodor (1991), and discussed in Block (1990), could be summarized as follows:

(11)

- (a) If there exists a conceptual link between a property A and a property B, then there cannot be a causal link between them.
- (b) There exists a conceptual link between a functional property and the properties mentioned in the functional role associated with it.

(c') There cannot exist a causal link between a functional property and the properties mentioned in the functional role associated with it.

The argument is obviously valid, and the conclusion (c') is obviously incompatible with our account of the causal powers of a functional property. Premise (b) also seems perfectly correct. It seems conceptually true that the instantiation of a functional property F requires the instantiation of a realizer of F. And it is conceptually true that the instantiation of a realizer of F involves the instantiation of a property causing the properties mentioned in the functional role of F (again this should be qualified in the case of more complex functional properties, but discussion of the simpler case will be enough for our purposes). Thus, for example, it is conceptually true that the instantiation by an entity X of solubility in water requires the instantiation by X of a property causing the dissolution of X in water under normal conditions. So in effect there seems to be a conceptual link between a functional property and the properties mentioned in its functional role. As for the premise (a) the main reason for it is the following. We all know of the limitations of a purely counterfactual analysis of causation. The sheer fact that a certain counterfactual dependence relation holds between two properties (or two instantiations thereof, strictly speaking), does not entitle us to infer a causal relationship between

them, or the causal efficacy of one of them with respect to the other. We have well known counterexamples to such an inference. For instance, there exists a relation of counterfactual dependence between the property of being dead as instantiated by Socrates and the property instantiated by Xantippa of becoming a widow, but we would not infer from that that there exists also a causal relation between these two instantiations. One plausible explanation as to why we do not license the inference in this case is that we know a priori that the corresponding counterfactuals involving these instantiations are true. We know this for conceptual reasons, and this seems to lead us to discard any causal connection between them.

If this reasoning were correct, then indeed premise (a) of our argument would be correct and we would have a rebuttal of our proposal. But there seems to be something wrong in this reasoning, since premise (a) would seem to be clearly wrong. For instance, to consider a simple example, in classical mechanics there exists a conceptual link between the force acting upon an object and its acceleration, and this fact does not prevent us from attributing also a causal connection between them. Of course, the causalist interpretation of the laws of classical mechanics, and the theory of the meaning of theoretical terms implicit in this counterexample, are open to objection, but the point at stake here is that they are not rendered incoherent to make our premise (a) false.

However, this rebuttal of the modal argument may be too hasty. There is a crucial distinction, one could argue, between the force-acceleration case and the functional properties-properties mentioned in the functional roles case. In the first case, the hypothesis of the causal connection between the force acting upon a body and the acceleration of that body has an explanatory role in classical mechanics. It helps to explain many things, including the covariance between these magnitudes that the second law of motion describes. However, in the second case, the hypothesis of the causal connection between a functional property and the properties mentioned in its functional role does not have any explanatory role. Nothing seems to be crucially explained by invoking such a connection, not even the modal covariance between these properties, which is explained by the conceptual link between them. As a general criterion, we tend to abandon hypotheses when there is no explanatory role left for them. For instance, we discard any causal connection between double effects, given that we can explain the modal covariance between them in terms of a common causal source and there is nothing which is crucially explained by such a causal hypothesis.

So then, we could reformulate our modal argument in a more cautious way by weakening our first premise as follows:

(12)

- (a') If there exists a conceptual link between a property A and a property B, then we have a prima facie reason against the causal efficacy of A with respect to B.
- (b) There exists a conceptual link between a functional property and the properties mentioned in the functional role associated with it.

(c'') We have a prima facie reason to discard any causal connection between a functional property and the properties mentioned in the functional role associated with it.

The reason in favour of (a') is that in being a conceptual link between A and B we can in principle explain the modal covariance between A and B, and in case there is no explanatory role left to the hypothesis of a causal connection between A and B our general criterion would push us toward abandoning such a hypothesis.

Admittedly, (c'') is not overtly incompatible with our proposal about the causal powers of functional properties, and hence this modified modal argument (12) is not a knock-down argument against it, but it has considerable force, and obliges us to argue for the explanatory role of the causal connection of functional properties with the properties mentioned in their functional roles, in order to secure our efforts to render plausible the claim of the Causal Autonomy of the Mental. In fact, the situation becomes even worse when we realize that, according to our proposal, all the properties which are allegedly causally influenced by a functional property are conceptually linked to it.

These reasons suffice, it seems to me, to look for other possible replies to the modal argument. Thus, some philosophers have argued that premise (a') is also false. I will end this work by examining this objection.

The main reason against (a') is contained in the following proposition:

- (13) A property causally related to another property may also appear conceptually connected to it if it is described in a proper way, in the same manner as an event causally connected to another may appear as conceptually linked to it when described in the appropriate way. (Block 1990; Segal, Sober 1991).

The argument against (12), as we see, proceeds by analogy with a famous observation for the case of events contained in Davidson (1963). Suppose an event A causes an event B, then we can describe A as "the cause of B". Under that description, A appears conceptually linked to B, but this constitutes no prima facie reason against the causal connection between A and B. On the contrary, the

existence of such a connection is what entitles us to describe A as "the cause of B". The argument, then, is that what goes for events is exactly the same as what goes for properties. And this seems undeniable. Suppose that property P is causally efficacious with respect to property Q, then we can describe P as "the property whose instantiation causes the instantiation of Q under such and such conditions". Now, under that description, P appears conceptually linked to Q, but this does not prevent us from claiming its causal efficacy with respect to Q. Again, the presence of such a causal connection entitles us to describe it the way we do.

The upshot of this reasoning is that properties, like events, are conceptually linked to other properties only under certain descriptions, whereas they are causally linked to other properties depending on the existing causal laws, regardless of the way in which the properties are described. This reasoning presents us with a way round the modal argument, but it does not show us whether we can avail ourselves of this route. Let me elaborate a little the difficulty I have in mind here.

We may distinguish between direct and indirect conceptual links between properties. For instance, in the Davidsonian example above involving two events, A appears conceptually linked to B only in an indirect way, that is, because A has a property, the property of being the cause of B, which is (the property, not the event) conceptually linked in a direct way with B. Likewise, in the example involving property P and property Q, P appears conceptually linked to Q only in an indirect way, that is, because P has a property, the property of being causally efficacious with respect to Q, which is conceptually linked to Q in a direct way. Hence, in this case we have two properties, property P and the property of being causally efficacious with respect to Q, which are conceptually linked to Q in a different way. The first is linked only in an indirect way, in virtue of having the second property which is linked in a direct way. The conceptual connection involving P depends clearly on the conceptual connection involving the property of being causally efficacious with respect to Q. And the latter connection holds only in virtue of the very nature of the property of being causally efficacious with respect to Q. These facts justify my terminology. Thus, I propose in general the following proposition:

- (14) A property P is conceptually connected to another property Q in an indirect way if and only if the explanation of that conceptual connection involves the fact that P has a property which is directly connected to Q. A property P is connected to another property Q in a direct way if and only if the explanation of such a connection lies in the very nature of P.

Armed with this distinction now we can formulate very clearly the difficulty I alluded to earlier in this reply to argument (12). The difficulty is that all that the Davidsonian reasoning shows is that the existence of an indirect conceptual link between two properties is not a *prima facie* reason against the causal connection between the two, but it shows nothing similar in the case where the conceptual link among the two properties is direct, not indirect. In fact, in our examples we are hardly tempted to attribute any causal connection between the property of being causally efficacious with respect to Q and Q itself, or between the property of being the cause of B and B itself. Now, if we conceive of a functional property F as the property consisting in having a property with the right functional role (as is, by the way, usually the case) then the conceptual link between the functional property and the properties mentioned in its causal role is direct, and in that case the Davidsonian considerations are of no help for dealing with the modal argument (12). To put it in other words, the Davidsonian considerations are not a good reply to this further modification of the modal argument:

(15)

- (a'') If there exists a direct conceptual link between a property A and a property B, then we have a *prima facie* reason against the causal efficacy of A with respect to B.
- (b'') There exists a direct conceptual link between a functional property and the properties mentioned in the functional role associated with it.

 (c'') We have a *prima facie* reason to discard any causal connection between a functional property and the properties mentioned in the functional role associated with it.

Therefore, Davidsonian considerations are of help for dealing with argument (15) only in case we can challenge premise (b'') in that argument, that is, only in case we offer a plausible metaphysics of functional properties according to which the conceptual link between them and the properties mentioned in their functional role is only indirect. We need to frame, thus, an account of functional properties that is more complex than is usually assumed. An account according to which a functional property F is construed as a property having the property of being associated with a certain functional role, but not as consisting in that property. Moreover, our account of functional properties should respect all the requirements and demands made explicit in the first section of this article in order to be useful for the defender of the Causal Autonomy of the Mental. This forces us to make our metaphysics of functional properties

more sophisticated in ways I am not at present capable of envisaging. I will close this discussion by examining a possibility and arguing why it is not acceptable.

The possibility in question is encapsulated in the following proposition:

- (16) A (multiply realized) functional property consists in the disjunction of its realizers. It is functional in the sense that it can be denoted by a description following this pattern: "the property consisting in the disjunction of all properties having a given functional role Rf".

For example, following the suggestion in (16) we should understand solubility in water as the property consisting in the disjunction of all realizers of solubility in water, that is, as the disjunction of all properties causing the dissolution of that which instantiates them when it is submerged in water. Dormitivity would be construed as the disjunction of all properties causing sleep, and so on.

Conceiving of functional properties as disjunctive properties in this way, means a functional property is not conceptually linked in a direct way with any of the properties mentioned in its functional role, but only in an indirect way. For example, solubility is a disjunction of physical properties and it is conceptually linked to dissolution in water only because it has the property that all the members of the disjunction cause dissolution in water in normal conditions, that is, the conceptual link is merely indirect.

What is the problem then with this metaphysics of functional properties? To appreciate the problem, recall that we started our investigation trying to attain a metaphysics capable of giving support to the commonsensical widely shared view of the Causal Autonomy of the Mental, that is, to Nonreductive Physicalism and the causal efficacy of the mental. Well, the problem with this last proposal is that it does not look as a nonreductivist proposal. What reason is there to suppose that a disjunction of physical properties is not itself a physical property? Notice that each disjunct conceptually entails the disjunction. Are we willing to say that the disjunctive property is a property of an entirely different kind from that of the members of the disjunction? On what grounds? The reductivist minded philosopher used to respond to the objection of the multiple realizability of special science properties precisely by identifying the multiply realized properties with the disjunction of their realizers. The standard objection to that move was that it incorrectly assumed that a disjunctive property would be causally efficacious if its disjuncts were (Fodor 1974). Now our metaphysical proposal, having wandered down a tortuous and somewhat contrived path, reaches the same conclusion and meets the reductiv-

ist claim. We cannot coherently blame the reductivist strategy and at the same time accept this proposal which, in the last resort, amounts to the same thing.

Notes

† Research for this work has been partially funded by the DGICYT of the Spanish Ministry of Education as part of the research projects PB95-0760 and PB98-0495-C08-07. I am indebted to José Antonio Díez, Manuel García-Carpintero, Terence Horgan and Manuel Liz for helpful discussions of the ideas contained herein, and to Marta Masergas and Michael Maudsley who carefully revised, and improved, the English text.

BIBLIOGRAPHY

- Block, N.: 1990, 'Can the mind change the world?', in G. Boolos (ed.): *Meaning and Method. Essays in Honor of Hilary Putnam*, Cambridge, Cambridge University Press, pp. 137-170.
- Davidson, D.: 1963, 'Actions, Reasons and Causes', *The Journal of Philosophy* 60, 685-700.
- Fodor, J.: 1974, 'Special Sciences, or The Disunity of Science as a Working Hypothesis', *Synthese* 28, 9-25.
- : 1991, 'A Modal Argument for Narrow Content', *The Journal of Philosophy* 88, 5-26.
- Kim, J.: 1984, 'Concepts of Supervenience', *Philosophy and Phenomenological Research* 45, 153-176.
- : 1992, 'Multiple Realization and the Metaphysics of Reduction', *Philosophy and Phenomenological Research* 52, 1-26.
- : 1993, 'The Nonreductivist's Troubles with Mental Causation', in Heil, Mele (eds.): *Mental Causation*, Oxford, Clarendon Press, pp. 189-210.
- : 1997, 'The Mind-Body Problem: Taking Stock after Forty Years', in J. Tomberlin (ed.): *Philosophical Perspectives, vol 11: Mind, Causation and World*, Cambridge, MA, Blackwell, pp. 185-207.
- : 1998, *Mind in a Physical World*, Cambridge, MA, The MIT Press.
- Lewis, D.: 1994, 'Lewis, David: Reduction of Mind', in Guttenplan (de.): *A Companion to the Philosophy of Mind*, Cambridge, MA, Blackwell, pp. 412-31.
- Putnam, H.: 1967, 'The Nature of Mental States', in Putnam (1975, pp. 429-440).
- : 1973, 'Philosophy and our mental life', in Putnam (1975, pp. 291-303).
- : 1975, *Mind, Language and Reality*, Cambridge, Cambridge University Press.
- Segal, G., Sober, E.: 1991, 'The Causal Efficacy of Content', *Philosophical Studies* 63, 1-30.
- Schiffer, S.: 1987, *Remnants of Meaning*, Cambridge, MA, The MIT Press.

David Pineda is Assistant professor of Philosophy at the University of Girona (Spain) since September, 1994. His main lines of research are mental causation, the mind-body problem and the metaphysics of intentional action. He has papers on these topics in *Enrahonar, Revista de Catalunya, Teorema* and *Philosophical Issues*. In December, 1998, he reads his Doctoral Dissertation titled *Problemas y perspectivas del Fisicismo* at the University of Barcelona (Spain).