PERIODICA POLYTECHNICA SER. SOC. MAN. SCI. VOL. 10, NO. 1, PP. 163–176 (2002)

CAUSAL EXPLANATION IN THE SPECIAL SCIENCES: TWO MODELS*

Tamás DEMETER

Budapest University of Technology and Economics
H–1521 Budapest, Hungary

Received: October 10, 2001

Abstract

This paper argues against the assimilation of two models of causation and causal explanation within the physicalist camp. One of them is Jaegwon Kim's model of supervenient causation; the other is Frank Jackson's and Philip Pettit's model of programme explanation. Assimilating moves have been made on both sides. This paper aims to show that these moves are flawed because differences between the models can easily be shown once the supervenience theses laying behind them are reconstructed. This initial difference is shown to entail several further discrepancies. At the end of this paper, we try to make sense of 'causal relevance,' on which the idea of a programme explanation rests, in order to prevent further attempts to assimilate the two models.

Keywords: supervenience, causation, programme explanation, reduction.

1. Introduction

An almost commonsensical and very much alive part of the legacy of the 'unity of science' movement is a hierarchical view on the arrangements of scientific disciplines. On this traditional picture, physics is at the bottom of this stratification starting from which other disciplines, through chemistry, biology, psychology, etc., build upon each other and culminate on the top in the sciences of society. This picture has immense intuitive force. First of all, the picture proceeds from the sciences of the simple to that of the complex. Thus it seemingly matches our initial mereological intuition that the increasing complexity of phenomena is a result of the combination of basic ingredients whatever they may be. Moreover, it also enables us to explain fruitful theoretical interactions between neighbour disciplines. It makes easy to explain why it is possible for psychology to be useful for sociology, and for neurology to be useful for psychology, and so forth. Originally, this picture was designed to serve certain reductionist intentions. As PUTNAM and OPPENHEIM explicate this model,[1] the sciences standing higher in this hierarchy are expected

*I am indebted to Alexander Bird, Jessica Brown, Gábor Forrai, Peter Lipton, Hugh Mellor, Bence Nánay for helpful comments and discussion. I also wish to thank Michael Bresalier for improving my English. While writing this paper I was supported by an Eötvös Fellowship. An earlier version of this paper was read to an audience at the University of Bristol.

[1]Cf. Putnam and Oppenheim, 1958.

to reduce to lower level ones. The procedure of theory reduction for them can take various routes. One may try to define and thereby reduce the terms of a science in terms of a reducing science, thus providing a unified vocabulary; or the laws of a science can reduce to the laws of some other discipline resulting in explanatory unity. But one need not be committed to reduction in order to endorse a hierarchical picture like this. One may refuse to accept the possibility of reduction for various reasons, while maintaining that this stratification is still an apt characterisation of the interdisciplinary relations. FODOR,[2] for instance, argues for the 'autonomy' of special sciences on the basis that there are law-like regularities in nature that are not reducible to the laws of more basic sciences, because the concepts employed in the formulation of these laws cannot be translated into one another and thus the causal taxonomies of the special sciences do not map unequivocally onto that of the 'more basic', physical sciences.

Along similar lines, physicalists are frequently divided into two camps, as reductivists and non-reductivists. Both positions raise questions in connection with the internal relations of this hierarchical structure, especially ones that concern the status of the special sciences, namely those that cannot be labelled without further ado as physical sciences. More specifically, the questions usually concern a satisfactory account of the nature of special-science properties, their relations to each other and to more fundamental properties, and of the explanatory role they can play. At the present stage of the controversies, there is a proliferation of theories that argue either for various versions of reductionist or anti-reductionist strategies in approaching these problems.[3] Due to the unsettled state of the art, unifying tendencies are to be welcomed. In his recent book Jaegwon KIM takes a step towards this direction, as he proposes to assimilate two models that explain the metaphysical status of special sciences. He argues that his model of *supervenient causation* (*SC*) is essentially identical to the model of *programme explanation* (*PE*) as it is put forward by Frank JACKSON and Philip PETTIT.[4] As KIM puts it: *PE* 'does not substantially differ from what I used to advocate as a way of handling mental causation – the model of 'supervenient causation'.'[5] Essentially the same position appears in JACKSON,[6] who agrees with KIM in approaching *PE* as a special version of *SC*. Later the paper will focus on KIM's argument for the identification mainly because JACKSON thinks their identity self-evident and not in need for argument. KIM thinks that the models share a commitment to epiphenomenalism at least as far as supervenient properties that are invoked in special science explanations are concerned. Thus supervenient causal relations are also epiphenomenal because real causal processes must be located among the properties of the supervenience base. And epiphenomena can have only informational relevance. They are causally impotent, though they can provide information about the underlying processes. According to his argument,

[2] Cf. Fodor, 1974 and also 1997.

[3] Cf. e.g. Fodor, 1997; Woodward, 2000.

[4] Cf. Jackson and Pettit, 1988, 1990, 1992.

[5] Kim, 1998: 74.

[6] Jackson, 1998b: 102.

there is nothing more to JACKSON and PETTIT's proposal than this same perspective.

This paper will evaluate KIM's judgement, and will show that KIM is mistaken in various respects. First, I provide a short introduction to the models at issue. Second, I try to reveal what could have led him and JACKSON astray in identifying the models. Then I draw some conclusions. I compare the two models, arguing that KIM is wrong to claim that the two models are identical because, despite superficial and possibly deceptive similarities, *SC* and *PE* prescribe different constraints on the domains they concern. To place my cards on the table, my argument should show that

1. the metaphysical background implicit in *PE* differs in important respects from KIM's *SC* model of explanation; and that
2. *SC* is entailed by a specific view on explanation argued for by KIM himself, which is incompatible with the core idea of *PE*.

Finally, I complete my argument by showing that *PE* may make sense as a form of causal explanation, without being identical with *SC*. Let me add, however, that in this paper I do not evaluate either *SC* or *PE*; I would only like to correct a mistaken view, and through this correction to show how to interpret *PE*.

Although the direct aim is to show that the two models are not identical, I note that this result also has positive relevance. The argument contributes to the clarification of what the models imply. The proper understanding of their implications has remarkable significance for the debates about the status of special sciences. JACKSON and PETTIT[7] formulated *PE* with the intention of solving some puzzles about the causal character of psychological explanations appealing to broad mental content. In their later work their proposal was extended so as to cover various kinds of structural, functional, disjunctive, etc. properties, thus making it suitable for accounting for the explanatory structure of the special sciences in general.[8] KIM's *SC* was designed to be a solution to similar problems in the same domains. Thus, if my argument is correct, we will see that what we actually have here are two rival theories, instead of having two, merely superficially different versions of the same suggestion.

2. The Models in a Nutshell

2.1. SC

In order to understand the idea of *SC*, first we need to understand the world in which such causal relations may obtain. This world is a microdeterministic world where the characteristics of wholes strongly supervene on the characteristics of their microconstituents. This means, in turn, that 'a general claim of macro-micro

[7] Jackson and Pettit, 1988.

[8] Jackson and Pettit, 1990, 1992.

supervenience then becomes the Democritean atomistic doctrine that the world is the way it is because the microworld is the way it is'.[9] If strong supervenience is combined with KIM's Democritean commitment to microdeterminism, then we arrive to the thesis of *mereological supervenience*.[10] One possible formulation goes as follows: For any macro property F of an individual, there is a set G of some combinations of microphysical properties of the parts of the individual such that it is nomologically necessary that any individual instantiating any member of set G instantiates F.

If macro-properties must be understood on the mereological picture, then they do not have independent causal powers. Causal potential is rooted at the macro levels in a mereological way in the microcomponents of individuals. Therefore there are no mysteriously emergent causal powers that could be irreducibly separated from the properties and relations that belong to the individual. This gives rise to the *causal inheritance principle*:[11] If F is instantiated by being realised by G then the causal powers associated with F must be identical with, or be a subset of G's causal powers. It follows that when F is realised on another occasion by another set of properties, then the causal power of this instance of F will be identical with that of its actual realiser. Therefore, as KIM himself also puts it,[12] *SC* is a kind of epiphenomenal causation. Thus the definition of *SC* is as follows (cf. KIM, 1984: 99; 1998: 74): x's having F is a supervenient cause of y's having F^* if and only if x has G and y has G^* that subveniently fix F and F^* respectively, and there is an objective causal connection between G and G^*. This means that the causal relation $F \rightarrow F^*$ holds only in virtue of there being a subvenient causal connection $G \rightarrow G^*$.

2.2. PE

The primary question that JACKSON and PETTIT seek to answer concerns the problem of the role played by broad psychological properties, i.e. ones that are, at least partly, rooted outside the mind. How is it possible for broad mental states to have a causal role in bringing about behaviour, how can they figure legitimately in causal explanations? As JACKSON and PETTIT argue in their 1988 paper, it is appropriate to distinguish between two kinds of properties invoked in causal explanations, and analogously, between two kinds of causal explanation that explain in virtue of these properties. *PE*s refer to properties which are not causally *efficacious* with respect to a given event, but causally *relevant* to it. This highlights that there are two distinct ways in which causal explanations can be formulated. One is process explanation, a traditional way that appeals to properties causally effective in a process. The

[9] Kim, 1998: 18.

[10] Cf. e.g. Kim, 1997: 192; and also Enç, 1995: 169.

[11] Cf. e.g. Kim, 1998: 54.

[12] Kim, 1984: 95ff; 1998: 74f.

second kind of causal explanation is a *PE*, which does not cite a causal history construed in terms of the causally efficacious process. Rather it appeals to functional, disjunctive, or relational properties. *PE*s give modal information in the sense that they cover a range of possible situations each of which could have produced the actual result and one of them actually did. The realisation of a higher-level property ensures that there is an efficacious property, which performs the causal work, while the higher-level property itself does not figure in the efficacious process.

3. What the Two Models Share

Prima facie, it may not be surprising that KIM and JACKSON think the two models identical. Both imply strong physicalist commitments, but – and this is what they overlook – they do so in different ways. The proponents of the two models may agree with the main points of PETTIT's summary of the claims that must be endorsed by those who identify themselves with the label 'physicalist'.[13] What PETTIT does in this paper is to explain the content of the central physicalist claim, namely that the empirical world, as CRANE and MELLOR put it, 'contains just what a true complete physics would say it contains'.[14] PETTIT breaks up this claim into two pairs of theses that must be accepted by those who endorse a physicalist metaphysics. The theses are as follows:[15]

1.a There are microphysical entities; and

1.b they constitute everything.

2.a There are microphysical regularities; and

2.b they govern everything.

So far *SC* and *PE* need not differ, KIM and JACKSON are apparently right in identifying the models. The agreement in the content of the two models' physicalism can easily suggest that what they think about causation and explanation can be equated. One source of the mistake of identification may result from this agreement as far as the background of the models is concerned. The real differences lie below this superficial isomorphism. In the following I will identify two main differences. The first is concerned with the commitment to the existence of a bottom physical level. I argue that *SC* posits a bottom physical level, while *PE* is independent of any commitment of this kind. Secondly, I argue that there is a substantial difference in the specific forms of supervenience theses which *SC* and *PE* are built upon.

[13] Pettit, 1993.

[14] Crane and Mellor, 1990: 186.

[15] Pettit, 1993: 214ff.

4. Metaphysical Differences

4.1. Levels: Bottom or Infinitely Descending

According to *PE*, the realisation of a higher-level, programming property ensures that there is an efficacious property, which performs the causal work, while the programming property does not figure in the efficacious process itself. If the programming/efficacious distinction is taken in the absolute sense, then efficacious properties – and, correspondingly, causal explanations in terms of efficacious properties – can be located presumably only on the fundamental physical level. However, for pragmatic reasons, JACKSON and PETTIT allow for a relativistic interpretation of the distinction.[16] As we are rarely in the position to identify the causally efficacious microphysical properties, 'most of the explanations we are ever likely to offer will be of the programme variety.'[17] But *PE*s presuppose the presence of an efficacious property, therefore we need to locate the efficacious properties relative to the programming properties. In this sense mental properties can be taken to be efficacious in relation to social properties; neurological properties in relation to mental properties; bio-chemical properties in relation to neurological properties, and so on. The distinction can always be drawn and re-drawn with an eye to the explanation we want to give. And this picture is consistent with an infinitely descending regress of levels: there is no need for a fundamental physical level in order to make sense of causal explanation.

KIM is not clear about the relation of his model to a bottom level. Sometimes, he is ready to acknowledge that 'there is a bottom level, one consisting of whatever microphysics is going to tell us are the most basic physical particles out of which all matter is composed,' and also admits that the multitiered picture 'carries the assumption that there is a bottom tier, a layer of entities that have no physically significant parts'.[18] Elsewhere he says that 'the layered model as such does of course not need to posit a bottom level; it is consistent with an indefinitely descending series of levels'.[19] Still, it does not take too much effort to reveal that his model entails a commitment to a bottom level.

Supervenience relations create a layered world where lower levels include the entities whose properties are thought to determine higher ones, and the causal relations on the subvenient levels also determine supervenient causal relations. Given that higher levels are completely determined by the micro levels, two individuals are indiscernible in relation to level L if and only if they are *micro-indiscernible*. KIM's definition is as follows: For the properties of any two individuals belonging to L, if they are indiscernible in relation to properties of their constituents at all levels lower than L, then they are indiscernible with respect to all properties at L.[20] Thus

[16] Jackson and Pettit, 1990: 116.

[17] Jackson and Pettit, 1988, 1990, 1992 provide a variety of examples.

[18] Kim, 1993a: 337; cf. also 1998: 15.

[19] Kim, 1998: 123n23.

[20] Cf. Kim, 1998: 17.

by definition, indiscernibility obtains only if there is a *complete* decomposition of the individuals into non-overlapping proper parts, and complete decomposition is possible just in case there is a bottom level on which decomposition can taken to be finished. This ensures that there is a level of causation where genuine causal relations obtain, and which supervenient causal relations are built upon. The question is then whether KIM needs this strong concept of indiscernibility, or a weaker one would do as well. Perhaps it would be enough to say that two individuals are indiscernible at L if there obtains indiscernibility at any level lower than L.[21] This may seem sufficient for indiscernibility at L, and also at levels higher than L, because if we take the indiscernible level as the supervenience base for L then the indiscernibility of the supervenience bases imply indiscernibility at every supervenient level. This seems to allow for getting along without the identification of the bottom level.

But this route is not open for KIM for he proposes the following argument elsewhere.[22] Higher-level properties are typically multiply realisable, i.e. higher-level properties may be realised by infinitely many different subvenient structures. As the causal power of these properties is dependent on their realisers, the causal power of two properties will be identical only if their realisers are identical as well. Differences in realisers may result in causal differences; therefore higher-level properties belong to causally heterogeneous kinds consisting of infinitely long disjunctions of realiser properties. This means that if the decomposition of two individuals stops on a level higher than the bottom one, the indiscernible common properties found may belong to a causally heterogeneous kind. They cannot guarantee the indiscernibility of properties supervenient on them because their causal power is entirely inherited from their actual realisers. If the realiser is different in the two individuals, then the outcome may turn out to be different at higher supervenient levels. Only full decomposition to the bottom level can ensure indiscernibility at all levels. Furthermore, if there were no bottom level then there would be no genuine causal relations in KIM's world: if every level is supervenient on another then causal relations are always supervenient causal relations. Consequently, while *PE* is neutral concerning the bottom level as the model is consistent both with and without it, *SC* is committed to the existence of a bottom level and does not make sense without it.

4.2. Supervenience: Global vs. Mereological

There is a more important point at which *SC* and *PE* are different. JACKSON and PETTIT created their model in order to solve the problem of broad psychological explanations. Psychological explanations refer to mental states; they explain behaviour, as it is commonly put, in virtue of beliefs and desires. Mental states are frequently defined functionally, in terms of the causal role they occupy. In this sense, being in a mental state means being in a certain neurological (realiser) state

[21] I owe this point to Peter Lipton.

[22] Kim, 1992.

under relational characterisation. However, since PUTNAM's famous argument,[23] it became common wisdom in philosophy that some mental states have broad content, since they are related to the state of affairs in the world in a way that has a significant role to play in determining their content. This entails that mental states cannot be internal neurological states under some specific relational characterisations, as functionalists argue. They must be more than that, since they are not exclusively in the head. Mental states do not supervene exclusively on the internal neurological structure, but on this internal physical structure *plus* some physical facts about the environment. Even if two persons are in the same physical state, the content of their thoughts, as one can conclude from PUTNAM's well-known Twin Earth example, may nevertheless be different, provided that there are differences in their environments. Now, the original challenge for JACKSON and PETTIT was to solve the puzzle arising from the tension of the functionalist view and the reality of broad content appealed to in psychological explanations. They did this by making a distinction between causal relevance and efficacy, claiming that mental states have the former but not the latter. The presence of a causally relevant mental state ensures, or programmes for, the presence of a causally efficacious property, which is in charge of the causal work.

For KIM the case to be solved is quite different. As is obvious, mereological supervenience does not allow for the existence of broad mental content in any philosophically interesting sense. On the mereological picture it is not possible for mental states, or special-science properties in general, to include properties of the external world, or even relational properties of individuals and their environment. For KIM,[24] the psychological one is supervenient exclusively on internal constitution, therefore what we need in, for example, psychological explanations can be given in terms of neurologically reducible narrow content properties that lack reference to any facts of the environment. Two persons being identical in all mereological levels lower than the mental one are identical psychologically – without respect to the differences in their surroundings. The proper causes of their behaviour can be fully located within their internal physical structure: if they are in the same physical state, their behaviour will be the same as well – given, of course, that they receive the same stimulus.

Therefore, on a closer scrutiny, *PE* and *SC* are found to be different against the background of the type of their authors' physicalist commitments. *PE* is designed to solve a problem that simply does not exist within KIM's physicalist framework. And as they are searching for solutions for only partially overlapping problems, they cannot provide substantially similar answers. Furthermore, there is a deeper issue here. Both being physicalist, the two models are committed to the supervenience of the mental on the physical. But the specific form and content of the supervenience theses on which they rely are quite different. Given the intention of accommodating broad content in a functionalist picture of the mind, *PE* is inconsistent with the mental being mereologically supervenient on the physical. The kind

[23] Putnam, 1975.

[24] Cf. also Kim, 1982

of supervenience that can properly be associated with *PE* can only take the form of a global supervenience thesis.[25] From this perspective, mental properties are supervenient on a *relation* between internal physical structure and certain facts about the environment. This means that mental properties are eventually supervenient both on extrinsic, relational properties and intrinsic ones, instead of being supervenient exclusively on intrinsic ones.

The roots of this difference can be understood in terms of the difference between the posits on which the two models are built. As we have seen above, the two models agree in the crucial physicalist theses on PETTIT's list, but they perhaps would not agree with the elaboration that PETTIT attaches to some of the theses. One of these is certainly the fourth remark added to (1.a),[26] which goes as follows: the physicalist 'may think of the microphysical realm in a non-atomist way; he may believe that certain relational, microphysical properties – apart from spatio-temporal properties – are in some way fundamental'. This is consistent with the global supervenience thesis that *PE* is built upon, but inconsistent with the Democritean atomism that KIM endorses and formulates in his mereological supervenience thesis. The case is quite straightforward: although mereological composition can allow, in the supervenience base, for relations between the properties of the individual and remain consistent with the local character of mereological supervenience, it is not open to include relational properties of the individual and its environment, only intrinsic ones. Otherwise the mereological thesis would cease to be local (and would not count as mereological either), and would turn into a global supervenience claim. This brings out the difference behind *PE* and *SC*: both are physicalist models but they envisage physicalism in different ways. *PE* entails a holistic, or non-atomistic, view of the world, while *SC* insists on a Democritean atomistic picture. Far from being identical, *PE* and *SC* are incompatible.

5. Constraints on Explanation

What do these differences amount to in relation to causal explanation? The *SC* model implies the principles of explanatory realism and explanatory exclusion, commitments not shared by *PE*. Behind this difference it is easy to recognise the initial divergence portrayed above concerning the supervenience theses on which the models are built. These are clearly consistent and follow inevitably providing that one accepts mereological supervenience as the basis of physicalism, but they are inconsistent with a physicalism based on global supervenience. I will introduce these two principles that KIM happily associates to *SC*, and I hope to show here that JACKSON and PETTIT cannot do the same with *PE*.

[25] Elsewhere Jackson (1998: 12) explicitly accepts a global supervenience thesis, which supports indirectly my present argument.

[26] Pettit, 1993: 215.

5.1. Explanatory Realism

KIM frequently tends to assimilate explanation and causation.[27] In his view they are two sides of the same coin: explanation is the subjective, causation is the objective side. This amounts to a difference in the way of thinking about them: explanation is the business of epistemology, while causation pertains to metaphysics. This difference, however, is not especially significant, because in order for an explanation to be true it must correspond to something objective in the world. As nothing can count as knowledge if not true, our knowledge essentially implies reference to objective reality, the knowledge that results from an explanation must conform to the relation of events. This attitude results in insisting on the following judgement on *PE*: if it intends to be a kind of causal explanation then it must necessarily be in accordance with explanatory realism: Proposition *C* can be an *explanans* of proposition *E* only in virtue of the fact that there is an appropriate (objective) causal relation between the events referred to in *C* and *E*.[28]

For KIM no explanation can claim to be causal if this condition does not apply. His reason is obvious: if an explanation is not adequate extensionally, if there is no appropriate causal connection between the events referred to in the *explanans* and the *explanandum*, then the explanation lacks the crucial causal component. And as KIM puts it, 'any weaker conception would merely cheapen the idea of causal explanation'.[29] If this is translated into JACKSON and PETTIT's idiom, KIM's 'appropriate causal connection' means a causally efficacious connection. Pretending that explanations that do not correspond with real causal connections between events are nevertheless causal explanations results in inevitably false explanations: explanatory realism can be properly read as a necessary condition for an explanation to be causal.

The very idea of *PE* is inconsistent with explanatory realism, provided that the distinction between causal efficacy and causal relevance is meaningful.[30] JACKSON and PETTIT clearly claim that, as they understand it, *PE* is a kind of causal explanation.[31] Causal, in spite of its reference to events which themselves are not causally efficacious. For them the causal character of *PE*s is retained by the fact that they appeal to connections that are causally relevant. What kind of causal role is causal relevance? Given the kind of supervenience thesis associated with *PE* it is clear that the events invoked in process explanations and in *PE*s, although entirely independent, are nevertheless different. This arises from the nature of global supervenience thesis. As we have seen above, psychological properties, as well as other special-science properties, are supervenient on internal physical constitution *and* some facts about the environment. This means that, according to *PE*, there is no property identity, or reducibility between physical and psychological properties.

[27] E.g. in Kim, 1987.
[28] Kim, 1987: 229f.
[29] Kim, 1998: 75.
[30] Child (1994: ch.6.) provides various reasons for the acceptance of this distinction.
[31] Both in Jackson and Pettit, 1988, 1990.

In virtue of global supervenience, properties referred to in *PE*s can be causally relevant, and thus causally explanatorily informative, without being causally efficacious. Programming properties ensure that there is a causally efficacious property at work because their supervenience base incorporates causally efficacious properties. Therefore, the acceptance of the programme model must entail the denial of explanatory realism. Thus the only question that arises concerns the concrete content of 'causal relevance'. We will try to clarify it under Section 6.

5.2. Explanatory Exclusion

Explanatory realism does not deny that explanations referring to causally inactive properties can have some sort of explanatory force. They can be usefully invoked in explanations, and their relevance is properly called 'informational relevance';[32] but they cannot be causally explanatory. Explanations appealing to causally inert, epiphenomenal properties can provide some information about the underlying causal processes, but they can be explanations only in a derivative sense. This is the role KIM thinks to be properly associated with *PE*. KIM's explanatory realism, due to the strong requirement to invoke causally efficacious properties in causal explanations, entails the principle of explanatory exclusion.[33] On this principle, there can be two (or even more) correct causal explanations of an event E only if either one of the two is incomplete, or one is dependent on the other. Suppose therefore that

(1) Explanation X explains event E by citing C as a cause, and also that
(2) Explanation X^* explains event E by citing C^* as a cause.

On the principle of explanatory exclusion the following must be true of the relation of these two explanations:

(a) $C = C^*$, i.e. they are nomic equivalents;
(b) C^* is reducible to C;
(c) C and C^* are co-ordinate causal factors, or
(c') C is a proper part of C^*;
(d) C and C^* are sequential causal factors;
(e) C and C^* are sufficient causes of E on their own right (a case of causal over-determination).

It follows, in each case, that E has only one complete and independent explanation, which is a proper causal explanation containing all the factors relevant in bringing about E. Other explanations of the same event have merely derivative significance, and against the background of the complete explanation do not improve the epistemic situation. As a corollary, the special sciences are not autonomous in

[32] Kim, 1998: 75.
[33] Kim, 1989: 257.

any relevant sense, because the explanations they give are wanting and reducible to some more fundamental level. Explanatory exclusion thus denies the possibility of there being any epistemic gain in special-science explanations over and above the complete explanation, which is provided by the basic sciences.

As *PE* provides an account of how the special sciences could promise more than a redundant and wanting story of what can be told completely on a lower level, *PE* is inconsistent with this principle as well. On *PE* it is possible to have two explanations of the same event, not excluding each other, without over-determining the event. JACKSON and PETTIT refer to PUTNAM's example to illuminate the case.[34] Imagine we have a one inch rigid cubical peg, a rigid surface and two holes on it, one of which is a circle one inch in diameter, the other is a one inch high square. It is now open to explain the failure of the peg to fit into the round hole in two ways. *X*) It is possible to cite microphysical laws and particles to explain, or to deduce from them, that the given microphysical structures allow only one trajectory by which it is possible for the peg to pass through the surface. *X**) It is also possible to refer to elementary geometrical facts, to the rigidity of the components, i.e. to higher-order properties, and thereby explaining the fact that the peg can only pass through the square hole. What is the relation between the two explanations? Explanation *X* involves the full story, and gives a description of the situation in terms of causally efficacious properties. Explanation X^*, however, cannot be reduced to it because X^*, and by appealing to geometry, provides information not involved in *X*; although the properties X^* refers to supervene on the properties *X* refers to: the presence of the structural features appealed to in X^* ensures, or programs for, the presence of a microphysical structure which is in fact causally efficacious in, and thus responsible for, producing the result.

6. Making Sense of Causal Relevance: Outlines of a Proposal

Given that *PE* is inconsistent with the principle of explanatory realism is it still worth calling it *causal* explanation? There are two ways of answering yes to this question, but only one of them seems promising. The easy and unsatisfactory answer is to argue that explanatory realism is a too strong requirement and should be replaced by a more liberal constraint on causal explanation. One can cite, for example, David LEWIS who claims that to explain causally means to provide causal information about how an event occurs.[35] As KIM himself is ready to acknowledge, reference to epiphenomena can provide information about the underlying causal processes, and thus *PE* would classify as a causal explanation. However, this answer falls short of our needs. On this account *SC* would be a causal explanation as good as *PE* is, and this would make the models isomorphic instead of showing their difference. Besides, KIM would not admit *SC* as a proper causal explanation, while JACKSON and PETTIT would insist that *PE* is in fact a proper one. Furthermore, this answer

[34] Jackson and Pettit, 1988: 394f.

[35] Lewis, 1986.

may seem to be ad hoc, as it doesn't make use of the idea of causal relevance that should ensure the causal character of the explanations given in virtue of *PE*. In order to retain this causal character we need to see what causal relevance consists in.

A property can be causally relevant if it affects the causal powers of its bearer, that is, if it makes a real causal difference. Otherwise it would be hard to imagine how a property could have any causal role to play. Now, how can we judge the causal power of a property? If a property is causally efficacious it means that it in fact brings about the effect. The problem is to distinguish a programming property from an efficacious one without losing causal powers in the former. Some may object that currently there is no account of causation that could catch the differences between causal explanations put in terms of relevant as opposed to efficacious properties.[36] Here I try to show how it is still possible to explain the kind of causation underlying *PE*. I propose an approach to causation in terms of raising the chances of effects. On this account it is not necessary for a cause to bring about an event in order for it to count as a cause, it is enough if it raises the chances of the effect. This simply means that the chance (ch) of an effect (E) is greater with a cause (C) than without it. To put it formally: $ch_C(E) > ch_{\sim C}(E)$.[37] Now, consider the *PE* model from this angle. We can be confident that programming and causally efficacious causes do not bring about their effects with the same chances. $ch_{C\,(\mathrm{prog})}(E) \neq ch_{\sim C\,(\mathrm{effic})}(E)$, for a given neurological state can cause behaviour with some chance, while the corresponding mental state can be cited as a cause with some different chance. This is because the supervenience base of the mental state incorporates, beside the neurological state, some other factors that effect its causal potentials. This entitles us to say that here we have two different causes and neither loses causal power to the other.

The core of this response is that the presence of programming properties makes real causal difference because they represent different causal potentials. This does not hold for mereologically supervenient properties, due to the causal inheritance principle (cf. Section 2.1). Still an objection may arise. One might point out that the approach I propose to distinguish between the two causal relations shows only a quantitative difference instead of a qualitative: $ch_{C\,(\mathrm{prog})}(E)$ is different from $ch_{C\,(\mathrm{effic})}(E)$ merely in degree and not in kind. On this reading the proposal is not sufficient to establish the difference between two kinds of causal explanation. The response to this criticism is as follows: the difference in degree is only the first step towards understanding their character. This difference merely reflects a deeper metaphysical gap between the nature of programming and efficacious properties. This gap, as we argued in Section 4.2, is due to the fact that the supervenience base properties of programming properties may include relational properties. Thus we arrived again to the initial difference of the supervenience theses underlying *PE* and *SC*. This ensures that programming and efficacious properties may be different, and that the causal potentials they have are also different in kind.

[36] Cf. Thalos, 1998.

[37] Mellor, 1995: 67ff.

References

[1] CHILD, W., *Causality, Interpretation and the Mind*, Clarendon, Oxford, 1994.
[2] CRANE, T. – MELLOR, D. H., There is No Question of Physicalism, *Mind*, **99** (1990), pp. 185–206.
[3] ENÇ, B., Nonreducible Supervenient Causation, in Savellos, E. E. and Yalçin, Ü. D. (eds.): *Supervenience: New Essays*, Cambridge University Press, Cambridge, England, 1995.
[4] FODOR, J., Special Sciences, or the Disunity of Science as a Working Hypothesis, *Synthese*, **28** (1974), pp. 97–115.
[5] FODOR, J., Special Sciences: Still Autonomous after All These Years, *Philosophical Perspectives*, **11** (1997), pp. 149–163.
[6] JACKSON, F., *From Metaphysics to Ethics: A Defense of Conceptual Analysis*, Clarendon, Oxford, 1998.
[7] JACKSON, F., Mental Causation, in *Mind, Method, and Conditionals*, Routledge, London, 1998b.
[8] JACKSON, F. – PETTIT, P., Functionalism and Broad Content, *Mind*, **97** (1988), pp. 381–400.
[9] JACKSON, F. – PETTIT, P., Programme Explanation: A General Perspective, *Analysis*, **50** (1990), pp. 107–117.
[10] JACKSON, F. – PETTIT, P., Structural Explanation in Social Theory, in Charles, D. and Lennon, K. (eds.): *Reduction, Explanation, and Realism*, Clarendon, Oxford, 1992.
[11] KIM, J., Psychophysical Supervenience, (1982) in Kim, 1993.
[12] KIM, J., Epiphenomenal and Supervenient Causation, (1984) in Kim, 1993.
[13] KIM, J., Explanatory Realism, Causal Realism, and Explanatory Exclusion, (1987) in Ruben, D-H. (ed.): *Explanation*, Oxford University Press, Oxford, 1993.
[14] KIM, J., Mechanism, Purpose, and Explanatory Exclusion, (1989) in Kim, 1993.
[15] KIM, J., Multiple Realization and the Metaphysics of Reduction, (1992) in Kim, 1993.
[16] KIM, J., *Supervenience and Mind: Selected Philosophical Essays*, Cambridge University Press, Cambridge, England, 1993.
[17] KIM, J., The Non-Reductivist's Troubles with Mental Causation, (1993a) in Kim, 1993.
[18] KIM, J., The Mind-Body Problem: Taking Stock after 40 Years, *Philosophical Perspectives*, **11** (1997), pp. 185–207.
[19] KIM, J., *Mind in a Physical World*, MIT Pess, Cambridge, Mass., 1998.
[20] LEWIS, D., Causal Explanation, in *Philosophical Papers*, vol. 2., Oxford University Press, Oxford, 1986.
[21] MELLOR, D. H., *The Facts of Causation*, Routledge, London, 1995.
[22] PETTIT, P., A Definition of Physicalism, *Analysis*, **53** (1993), pp. 213–223.
[23] THALOS, M., A Modest Proposal for Interpreting Structural Explanations, *British Journal for the Philosophy of Science*, **49** (1998), pp. 279–295.
[24] WOODWARD, J., Explanation and Invariance in the Special Sciences, *British Journal for the Philosophy of Science*, **51** (2000), pp. 197–254.