

Article

Causal Relevance and Heterogeneity of Program Explanations in the Face of Explanatory Exclusion

Wilson Cooper

In everyday causal explanations of human behaviour, known generally as 'folk psychology,' the causal powers of the mental seem to be taken for granted. Mental properties such as perceptions, beliefs, and desires, are all called upon in causal explanations of events that are deemed intentional. Jaegwon Kim's exclusion principle has led him to deny mental properties causal efficacy unless they are metaphysically reduced to physical properties, but what of their causal relevance? By giving up the assumption of causally efficacious mental properties, has Kim put into question the explanatory value of explanations with mental descriptions? In other words, if a lower order neurological causal explanation involving a causally efficacious property is at hand, does it make the higher order mental explanation irrelevant and therefore redundant? If we are to save the explanatory importance of higher order predicates, and thus the causal explanations of the special sciences and folk psychology, we need an account of how such properties can be relevant as opposed to irrelevant in causal explanations, even though they may not be causally efficacious. Frank Jackson's and Philip Pettit's notion of program explanation tries to do just this.

There are two ways to read Kim's causal exclusion principle, metaphysically as causal exclusion, or epistemologically as causal explanatory exclusion. The explanatory exclusion principle hinders the acceptance of two causal explanations for a single effect unless an acceptable relation exists between the two purported causes. In this paper I will concentrate on the epistemological version of the principle, including Kim's acceptable relations, and contrast it with a survey of Jackson's and Pettit's program explanations. My aim is to discover whether program explanations expound an acceptable relation already covered by Kim or present an additional way of understanding how two causal explanations of the same phenomenon can be accepted.

Jackson and Pettit say: "The notion of a programming property does not just explain how an inefficacious property can be relevant to the causation of an event. It also shows how a program explanation can have a significance that remains in the presence of an explanation invoking the corresponding

efficacious property.”¹ The idea is that even though a causally efficacious property features in a causal explanation of an event, there can still be another explanation citing a distinct causally inefficacious property as the cause of the same event which has an explanatory force not possessed by the explanation citing the efficacious property. Such force is gained by the extra information provided in comparison to an explanation that cites the causally efficacious property. I will outline the notion of program explanation and investigate whether program explanations constitute a challenge to Kim’s explanatory exclusion principle by offering an account of the relation between higher and lower order properties not considered by Kim. As a result of this investigation, the homogeneity of Jackson’s and Pettit’s examples of program explanations will also be scrutinised. Further, I will consider Kim’s charge that program explanations are not causal explanations and thereby do not pose any explanatory competition to explanations citing causally efficacious properties.²

What is a Program Explanation?

In their paper “Program Explanation: A General Perspective,” Jackson and Pettit consider four plausible assumptions regarding properties that are involved in causal explanations; a problem these assumptions can be seen to generate; and then go on to propose a solution to the alleged problem.³ This problem can be seen to arise in connection with the situation mental properties, as higher-order properties, are left in by Kim’s causal exclusion thesis, *viz.* they are either reducible to their realisers or seen as causally idle. The assumptions are:

- (1) Causal explanations direct us to causally relevant properties as opposed to causally irrelevant properties.
- (2) One way a property can be causally relevant is to be causally efficacious.
- (3) A property F is not causally efficacious in the production of an effect if three conditions are all satisfied together:

Put briefly, these conditions are: (a) that there is a distinct property G, and F is efficacious in producing an effect *e* only if G is. (b) F and G are not sequential causal factors in producing *e*. (c) F and G do not jointly produce *e*.

¹ Frank Jackson and Philip Pettit, “Program Explanation: A General Perspective,” in *Analysis*, 50 (1990), 116.

² Jaegwon Kim, *Mind in a Physical World* (London: Massachusetts Institute of Technology, 1998), 75.

³ Jackson and Pettit, “Program Explanation: A General Perspective.”

There are a number of examples of properties that meet the three conditions above. One example used is a closed container that breaks because either its contents are at boiling temperature or there are certain molecules striking the container at a certain molecular bond with enough force to break it, etc. The boiling/temperature example is designed to illustrate a case where a property *F*, which is considered to constitute a causal explanation of an event, is inefficacious in accordance with the conditions of assumption (3). When a causal explanation is sought as to why a closed container holding boiling water breaks, the property of temperature is a commonplace answer. If we consider the temperature property as property *F*, we can see how the example works. In accordance with condition (a), the property *F* is only causally efficacious if another property *G* is also efficacious, in this case, property *G* will be the collision of a molecule at a certain point on the container. Condition (b) is satisfied because the temperature and molecular properties are not sequential causal factors in producing the breaking as either property can be instantiated prior to the other, or they can occur simultaneously. Finally, condition (c) is met since the two properties do not jointly combine to produce the effect, if the molecule hits the closed flask at the same point with the same momentum, the container can break despite the contents not boiling. The three assumptions outlined have some radical consequences when combined with the fourth assumption:

- (4) The only causally relevant properties in producing an effect *e* are causally efficacious properties.

The implications of the four assumptions are far reaching, for if they are all held together, special science explanations are not causal explanations at all and the properties they cite are not causally relevant to any effect.

Other examples Jackson and Pettit use include a vase breaking because of either its fragility or the molecular structure that makes the vase fragile. A square peg's not fitting into a round hole with the same diameter as the side of the square because of either its squareness or an impenetrable overlap. An elevator failing to move because either more than ten people entered or twenty people entered it. And a symphony conductor getting annoyed at someone from the audience coughing which can be explained by either someone from the audience coughing, or a particular member of the audience coughing, Fred for instance. I will generally keep to one example, the closed container with boiling contents, to make my argument easier to follow. However, there will be a need to highlight some of the other examples as we progress.

That the dire consequences from above are entailed by the four assumptions can be seen by the following inferences: For any property *F* in an explanation of a special science or folk psychology, there will be a different property *G* such that in the domain of a lower-level science, *F* and *G* meet conditions (a), (b), and (c) of assumption (3), making property *F* inefficacious. By assumption (4), property *F* is rendered irrelevant to the effect it is called upon to explain. Thus, by assumption (1) property *F* should not be included in

a causal explanation of that effect. Jackson's and Pettit's project is to show that assumption four is questionable and that being causally inefficacious does not entail being causally irrelevant.

The boiling/temperature example is a situation where a property, *viz.* the property of boiling, is inefficacious in bringing about the effect of breaking the glass. Intuitively, the fact that the contents are boiling is relevant for the breaking and this is supposed to leave us with the feeling that there are properties that are both relevant and inefficacious. A certain kind of explanation involves such properties, *viz.* program explanations, which invoke a "property-instance [that] does not figure in the productive process leading to the event but it more or less ensures that a property-instance which is required for that process does figure."⁴ The corresponding explanation that features the causally efficacious property is referred to as a 'process' explanation. Now the question turns to why a program explanation retains causal relevance when it is fully acknowledged that the property involved has no causal efficacy and a process explanation is at hand.

A program explanation can be significant, even when we have access to a lower level process explanation citing a causally efficacious property, because the process explanation does not contain all the information that a program explanation makes available. There are two kinds of information provided by the program explanation that are not conveyed by the process explanation. First, it is argued by Jackson and Pettit that a program explanation provides modal information, which a process explanation cannot supply; how things might have been rather than how things actually went, the information that the same effect would have ensued even if similar but different process properties had been instantiated.⁵ Second, there is the information that the program property, the boiling contents of the container for example, is instantiated. The fact that the water is boiling, carries with it additional information, *viz.* that there will be some suitably situated molecule that has the required momentum to break the glass. Even if property G, say molecule *x* colliding with molecular bond *b*, was not instantiated, it is "probable to a point approaching certainty" that another property, G₁, say molecule *y* colliding with molecular bond *j*, will be instantiated and the breaking will take place regardless of which molecule actually does the causal work.⁶

Are Program Explanations a New Relation Not Considered by Kim?

I now want to look more closely at the relationship between Kim's explanatory exclusion and Jackson's and Pettit's program explanations. The general principle of explanatory exclusion says that there cannot be two

⁴ *Ibid.*, 114

⁵ *Ibid.*, 117.

⁶ *Ibid.*, 115.

complete and independent causal explanations of a single event.⁷ Concerning a situation where we have two explanations of a single event, Kim states the explanatory exclusion principle as: “*No one may accept both explanations unless one has an appropriate account of how they are related to each other.*”⁸

In “Mechanism, Purpose, and Explanatory Exclusion,”⁹ Kim outlines five cases of how two explanations A and B that cite F and G respectively as a cause of event *e* can be related:¹⁰

- (1) F and G can be identical.
- (2) F is distinct from G and F is dependent on G, either by supervenience or realisation in such a way that F is reducible to G in some clear sense, although this will preclude identifying F with G. This case precludes the independence of F and G and thus, the independence of explanations A and B.
- (3) F and G are two independent but insufficient causal factors for an event, one may be more salient in a certain context but still incomplete. An addendum to (3) is:
 - (3a) That F is a proper part of G.
- (4) F and G are two links in a single causal chain, thus making one causally dependent on the other.
- (5) F and G are causally independent and sufficient for *e*, i.e. a case of genuine overdetermination. However, in most cases of overdetermination, an explanation of *e* would need to include both F and G to be a complete explanation.¹¹

In the example of explaining why the closed container of boiling water breaks, it would seem that we are faced with two explanations of the single event of the breakage, a program explanation, and a process explanation. We can say the program explanation which refers to the inefficacious property F, i.e. the temperature of the boiling water, is explanation A, and the process explanation citing the causally efficacious property G, i.e. a certain water molecule striking a particular molecular bond, is explanation B. Now, since the general principle states that both explanations cannot coexist if both explanations are considered independent and complete, the explanatory

⁷ Jaegwon Kim, “Mechanism, Purpose, and Explanatory Exclusion,” in *Supervenience and Mind* (Cambridge University Press, 1993), 250.

⁸ *Ibid.*, 257. Emphasis in original.

⁹ Kim, “Mechanism, Purpose, and Explanatory Exclusion.”

¹⁰ Kim’s original notation has been altered to remain consistent with Jackson’s and Pettit’s.

¹¹ Kim, “Mechanism, Purpose, and Explanatory Exclusion,” 250-252.

100 CAUSAL RELEVANCE

exclusion principle guides us to the conclusion that one is excluded unless the two explanations are related in one of the five ways outlined above. Since Jackson and Pettit rule out the causal efficacy of the higher order property F because property G is the causally efficacious property causing the container to break, it would seem that explanation A that cites property F is a prime candidate for exclusion. Note, however, that program explanation A provides more information than process explanation B alone, which suggests that the exclusion of a program explanation might be an epistemically implausible choice. Since Jackson and Pettit claim the causal relevance program properties exhibit is not derived from being efficacious but from the extra information they contain, it is also plausible that neither explanation is complete. We are now presented with two alternatives as to how explanations citing program properties are related to the explanatory exclusion principle. Either the two explanations can be accommodated by one of the cases outlined above, or program explanations are a further case of the relation that may exist between two explanations of a single event.

I will examine two of the cases put forward by Kim with regard to the examples of program properties proposed by Jackson and Pettit, and argue that Kim can account for the relevance of program properties. There are two of Kim's cases that seem, at first glance, to account for the relevance of program properties if we continue to allow that both A and B are causal explanations: case (2), F is distinct but dependent on G, or case (3a), G is a proper part of F. After dealing with these issues, I will consider the third possibility that program explanations are not causal explanations and, therefore, cannot be considered as competing with process explanations in any *causally* relevant sense.

Supervenience and Mereological Relations

Case (2) states that property F is distinct from G, but that F is nomologically reducible to, or supervenient on G in a way that allows for the higher-level causal relation to be dependent on the lower-level causal relation. I will make some suggestions as to why case (2) can explain the relevance of program explanations, but also that some features of the example in use might suggest that case (3a) is a more appropriate relation. I will take this consequence as support for claiming that the different examples offered by Jackson and Pettit will fall under different possibilities. First, though, I will demonstrate that there are a number of features of program explanations, and the use of them by Jackson and Pettit for supporting explanatory ecumenism in their paper "In Defence of Explanatory Ecumenism"¹² which generally point to the properties they cite being supervenient properties. In that paper, Jackson and Pettit make explicit their endorsement of "causal fundamentalism," a position that involves the "doctrine of supervenient

¹² Frank Jackson and Philip Pettit, "In Defence of Explanatory Ecumenism," in *Economics and Philosophy*, 8 (1992).

determination by the micro of the macro.”¹³ In addition, there are fundamental similarities that exist between program explanations and Kim’s solution to the problem that the causal exclusion principle generalises to all special science properties.¹⁴ Since, on Kim’s view, the supervenience relation is retained between inefficacious properties and their correlated efficacious properties, these similarities lend more support to considering the relationship between properties contained in program explanations and process explanations as one of supervenience.

To begin with, I will highlight some similarities that exist between the notions of program explanations and Kim’s proposed solution to the problem that the exclusion principle generalises to the special sciences. Kim accepts the idea that supervenient properties within a single level, which he calls higher-order properties,¹⁵ are inefficacious. According to Kim, all the properties of a single object are on the level of which the object is characteristic and particular properties of objects distinctive of levels are referred to as higher and lower *order* properties on his model.¹⁶ If we take the water contained in a closed flask as a single object, then the properties of that object, water, described at different levels will line up precisely with Kim’s hierarchy of orders, and in fact Jackson and Pettit use the terminology of ‘orders’ in their essay “Program Explanation: A General Perspective.” On this view, the molecular properties of the water can be straightforwardly considered as *subvenient* properties of the supervenient property of temperature.

Another similarity existing between program explanations and Kim’s solution is the point acknowledged by both proposals that inefficacious properties do provide some kind of information. For Kim “second order designators come in handy when we are not able or willing to name the properties we have in mind by the use of canonical first-order designators.”¹⁷ Besides coming in handy, Kim also thinks higher order properties, “serve important conceptual and epistemic needs.”¹⁸

I will proceed by examining Jackson and Pettit’s comments on ‘causal fundamentalism’ and the accompanying view that takes “causal relations at higher scientific levels as superveniently dependent on the causal relations and regularities that obtain at lower levels.”¹⁹ In our example where two explanations are given for the container breaking, we have the macro-property of temperature and the micro-property of molecular collision. As mentioned above, Jackson and Pettit endorse the view that macro properties and the causal relations that obtain between them are determined by, and supervene on

¹³ *Ibid.*, 4-5.

¹⁴ Ned Block, “Do Causal Powers Drain Away?,” in *Philosophy of Phenomenological Research*, 67 (2003); Jaegwon Kim, *Physicalism or Something Near Enough* (Princeton: Princeton University Press, 2005).

¹⁵ Kim, *Mind in a Physical World*, 80-87.

¹⁶ *Ibid.*, 82.

¹⁷ *Ibid.*, 104.

¹⁸ *Ibid.*, 110.

¹⁹ Jackson and Pettit, “In Defence of Explanatory Ecumenism,” 7.

102 CAUSAL RELEVANCE

micro properties and their causal relations. However, to hold supervenience as a general view of how micro and macro properties of the world are related does not entail that there is a supervenience relation holding between any particular pair of properties cited by explanations A and B. As we will see, this may not be the case.

In another example, two explanations of why an elevator failed to move are given: one explanation is that twenty people entered, the other explanation is that more than ten people entered.²⁰ Here the inefficacious 'more than ten' can be seen as supervening on the efficacious 'twenty'. For the program property, 'more than ten,' has a subvenient process property such that anything possessing the process property, i.e. twenty people, will also possess the program property, i.e. more than ten people. Furthermore, since the program property is determined by the process property, but not conversely, it seems appropriate to view the program property as dependent on the process property. Supervenient properties are generally thought to be instantiated synchronously. This may be a contentious issue with the elevator example on some interpretations, but considered as the causal antecedent of a single event, i.e. the failure of the elevator to move, they can be easily viewed as synchronous. In the case of the elevator example, these points strengthen the argument that Kim's explanatory exclusion principle allows for the explanatory relevance of the program explanation by virtue of the existence of a supervenience relation. Other examples, however, are not so easily accounted for by supervenience.

To use the example of the glass flask breaking due to the temperature of its contents again, the efficacious property is a suitably situated molecule striking/vibrating at a particular point on the flask with enough momentum to cause the breaking. The increase in kinetic energy of one molecule can conceivably be enough to increase the mean total kinetic energy (MTKE) of the whole of the contents, if there is not a drop in the kinetic energy of other molecules, and an increase in temperature is dependent on, and synchronous with, any increase of MTKE. However, they are not properties of the same object. Jackson and Pettit will reject that the two explanations are related by supervenience since the property of being a certain temperature does not supervene on the single molecule cited as being causally efficacious. It is possible that the temperature can be different without *the* single efficacious molecule being any different, and this situation precludes the supervenience relation holding between the two properties.

If this line were to be used in resisting the implication of supervenience as a possible relation between the properties of program and process explanations, it is still conceivable that the relation is a part/whole relation, as Kim's case (3a) points out. For when we are investigating the properties of a container of water, where the water is considered as a single object, at the molecular level, a discrete molecule with a particular kinetic

²⁰ Frank Jackson and Philip Pettit, "Causation in the Philosophy of Mind," in *Philosophy and Phenomenological Research*, 50 (1990b), 205-206.

energy can be construed as a proper part of the sample of water. In addition, the property of having a certain kinetic energy can be understood as being a part of the statistical mean, a property of the whole, in a straightforward sense. This possibility may help explain why explanations A and B provide distinct information, for if it is a part of the whole that is causally efficacious, the whole remains relevant by having the efficacious part in question. Being a whole of a particular kind will *ensure* the possession of parts of a particular kind.²¹

Another example which can possibly be seen as related by the part/whole relation is the example of a symphony conductor turning around in annoyance because of either, someone's coughing, or a particular person, Fred's coughing.²² In this case, the property of someone coughing can be seen as a property of the audience. If the audience is taken as the object, and thus the whole, then it is quite natural to suppose that Fred is a proper part of the audience. Again, the relevance of explaining the conductor's annoyance by a property of the whole, 'someone's coughing,' is retained by the fact that the causally efficacious property, 'Fred's coughing,' is a proper part of that whole. It could be claimed that a property of the part, i.e. Fred's coughing, is not a property of the whole in the sense that an explanation would invoke the audience as a whole of which a part, Fred, had the property of coughing. However, the example can still be accommodated by supervenience, since an explanation citing someone's coughing would be dependent on Fred's coughing. For in any possible world in which Fred coughs, someone coughs.²³

We can see from the discussion so far that the cases of program explanations considered do not present examples of a further relation in understanding how two causal explanations of a phenomenon can be accepted. All the examples can be accommodated quite accurately by the instances previously outlined by Kim. Therefore, the appropriate conclusion to draw from the analysis is that program explanations are consistent with Kim's principle of causal explanatory exclusion and offer nothing new in exemplifying the importance of higher-order causal explanations.

Are Program Explanations Correctly Conceived of as Causal Explanations?

I will now consider a third possibility. Kim states: "that program explanations, whatever their explanatory value, cannot be *causal* explanations . . . a causal explanation of an event that invokes another as its cause can be a

²¹ I have examined Jackson and Pettit's case here with the notion of causation assumed as originally presented and used Kim's framework to accommodate that claim. However, recent work into the contrastive nature of causal relations might reveal that there are two separate effects depending on the specificity of the effect under investigation. This would be another way in which explanatory exclusion could be avoided. I would like to thank an anonymous referee for directing my attention to this possibility.

²² Frank, Jackson and Philip Pettit, "Functionalism and Broad Content," in *Mind*, 97 (1988), 394 and "In Defence of Explanatory Ecumenism," 9.

²³ Again, I would like to thank an anonymous referee for bringing this point to my attention.

correct explanation only if the putative cause really is a cause of the event to be explained.”²⁴ The most natural way to read this admonishment of program explanations is in the metaphysical sense, for it is the inefficaciousness of program properties that the emphasis is placed on. Moreover, in *Mind in a Physical World*, Kim is focused on the metaphysical issues of mental causation rather than explanatory relations. The question I will concentrate on here, however, is whether there is any justification for a repudiation of the causal relevance of higher order properties in causal explanations.

The assumption Jackson and Pettit resist says: “The only way for a property to be causally relevant to the production of an effect is by being causally efficacious in the process of production.”²⁵ Again, we can discount the possibility of the causal relevance in this passage referring to metaphysical relevance since Jackson and Pettit state: “our causal fundamentalism reveals that [the higher level] property is causally related to the result only in virtue of causal relations at a lower, smaller grained level: a level that is not causally effected, of course, by what happens higher up.”²⁶ So, the information a program explanation is providing is for the benefit of the epistemically limited situation we find ourselves in. As discussed above, such relevance may be derived from a relationship between properties already accounted for by Kim’s explanatory exclusion principle. However, Jackson and Pettit accept that one of the examples used by them may not have persisting relevance when the corresponding process explanation is on hand.

The example of fragility is the case in point. In this example, we have two explanations of why a vase breaks on impact when dropped, either because it is fragile or because of its molecular structure.²⁷ The problem with this example is that there seems to be some avenue to a conceptual reduction of the two properties, “anyone who had access to the [molecular] account would have all the significant information at his disposal which is offered by the fragility explanation.”²⁸ This, I think, again gives support to the heterogeneous nature of Jackson and Pettit’s examples; if a conceptual reduction of some kind is possible on a particular occasion then explanatory exclusion will effectively remove the program explanation in the same way that metaphysical reduction will remove the higher-level cause.

Another possibility that may induce Kim into denying causal relevance to program explanations is the modal aspirations that program explanations lay claim to. In having a process explanation of an event we have a story about the actual cause of that event, and for the particular event in question, there is nothing more that we need to know to understand why it was caused. Viewed in this way, a program explanation will have no relevance to the specific event under scrutiny. The modal relevance of a program explanation has to do with

²⁴ Kim, *Mind in a Physical World*, 75.

²⁵ Jackson and Pettit, “Program Explanation: A General Perspective,” 111.

²⁶ Jackson and Pettit, “In Defence of Explanatory Ecumenism,” 10.

²⁷ Jackson and Pettit, “Functionalism and Broad Content,” 395; “Program Explanation: A General Perspective,” 109; “In Defence of Explanatory Ecumenism,” 10.

²⁸ Jackson and Pettit, “Program Explanation: A General Perspective,” 113.

future events with *similar* characteristics to that of the original event, which the process explanation has already causally explained. By “being made aware of the boiling water explanation, we learn something new: we learn that in more or less all possible worlds where the relevant causal process is characterised by involving boiling water, the process will lead to the flask cracking.”²⁹ The question is: in what sense is such pragmatically useful information, information used to predict future causal relationships, said to have ‘causal’ relevance? The information certainly gives its possessor the ability to predict that certain things will happen by pointing to correlated inefficacious properties. However, this accomplishment is achieved because of the epistemic accessibility to the inefficacious property. So, if it can be established that the program property is always correlated with some efficacious process property, then that would establish the default causal relevance of the program property.

However, the relevance does not look to be causal; it is more naturally construed as explanatory relevance. When we combine Jackson and Pettit’s acknowledgement of program properties as inefficacious with Kim’s criterion of realism regarding causal explanations, it is natural that Kim would deny both that program explanations are not causal explanations and that it is not causal relevance that program explanations possess. On the accounts of both sides, program explanations are relevant to arriving at causal explanations by serving as guides that point to the causally relevant properties at stake, and as Kim concedes: “no one should legislate what counts and what doesn’t count as explanation.”³⁰

The points raised here have relevance in connecting this debate to more contemporary developments in philosophy of mind, and in particular mental causation. Kim’s exclusion principle has been a catalyst for an increased interest in specifically formulating what the relationship is that exists between mental and physical properties, while avoiding the principle’s negative implications. This has led to concentrated research in the areas of causation and realisation that aims at individuating properties by their modal features.³¹ Some recent arguments indicate that the conception of causation Kim is assuming might be inadequate and as such may itself undermine the exclusion argument.³² So, increasingly, not only the relevance, but the causal autonomy of higher-order properties in causal explanations is gaining in legitimacy. These issues turn on some intricate details that are continuing to be contested in the literature and due to space limitations, cannot be included here, so an argument for a direct connection will have to wait. However, the notions canvassed here

²⁹ Jackson and Pettit, “In Defence of Explanatory Ecumenism,” 15.

³⁰ Kim, *Mind in a Physical World*, 76.

³¹ Judea Pearl, *Causality: Models, Reasoning, and Inference* (New York: Cambridge University Press, 2000); Sydney Shoemaker, *Physical Realization* (New York: Oxford University Press, 2007); James Woodward, *Making Things Happen: A Theory of Causal Explanation* (New York: Oxford University Press, 2003).

³² Peter Menzies, “The Exclusion Problem, The Determination Relation, and Contrastive Causation,” in *Being Reduced – New Essays on Reductive Explanation and Special Science Causation*, ed. by Jakob Hohwy and Jesper Kallestrup (Oxford: Oxford University Press, forthcoming).

and the concerns they highlight can be seen to have inspired, at least indirectly, some recent developments in the area of mental causation.

The Causal Relevance of Mental Properties

Do program explanations deliver an account of higher order properties that will vindicate the causal relevance of *mental properties* in causal explanations? Different examples of program explanations seem to present different possibilities as to how they are related to their correlated process explanations. I have attempted to show how some examples can be understood as displaying a mereological relation, and that this relation can justify the explanatory relevance of some program properties. Other examples represent a relation that is more consistent with the supervenience, or realisation relation, and within these examples, such relevance will depend on whether the supervenient properties are reducible to their base properties or only necessarily correlated with them. Higher-order properties that face conceptual reduction, like the example of fragility, would seem to be made irrelevant by explanatory exclusion, but what about mental properties?

Taken as functional and thus, higher-order properties, Kim claims mental properties are deprived of an independent causal efficacy due to their reducibility.³³ When contrasted with examples of program properties, they are meant to remain causally relevant because of the extra causal information they convey over and above the information contained in accompanying process explanations. We have seen though, that the relevance kept by different examples of program explanations can be retained for different reasons. A mental property that is supervenient on the physical in much the same way as 'more than ten' is supervenient on 'twenty,' will retain relevance both for Kim, and for Jackson and Pettit by providing *useful* information. Kim will question whether it is *causal* relevance. Construed as properties of a whole, a mental property may retain causal relevance in the same way that the example of boiling water retains its relevance by having a suitably situated vibrating H₂O molecule as a proper part. Some mental properties may be reducible to underlying subvenient properties, like the example of fragility, thus excluding the relevance of such properties in intentional explanations.

The plausibility of whether all mental properties are conceptually reducible to physical properties is something that even Kim denies. For Kim, mental properties that have intrinsic qualities are not physically reducible because they cannot be functionalised, and Kim believes this is true of qualia.³⁴ On Kim's account, however, this also leaves such properties as causally inefficacious and therefore causally irrelevant. But there are other reasons for denying that all mental properties are conceptually reducible to the physical. One is Donald Davidson's argument for the anomalous nature of events with

³³ Jaegwon Kim, *Physicalism or Something Near Enough* (Princeton: Princeton University Press, 2005), 165; *Mind in a Physical World*, 118.

³⁴ Kim, *Physicalism or Something Near Enough*, 173.

mental descriptions.³⁵ Davidson argues that mental and physical terms belong to incommensurable theoretical frameworks and because they are each governed by disparate commitments, any reductive attempt would dismiss those commitments and result in a categorical change in the subject. Another reason for denying the conceptual reduction of mental properties to the physical is multiple realisation. It is claimed that since mental properties can be realised by a number of different physical realisers, and those realisers themselves are distinct, then the multiply realised mental property must be distinct from the physical. Furthermore, there is reason to believe that the modal properties of mental property instances are different to the modal properties of the physical properties that realise them. Establishing that mental properties are distinct from physical properties is one possible way that could lead to the causal relevance of those properties. The distinctness of mental properties would then need to be supplemented with the conviction that they are the more plausible candidates, compared to competing physical properties, for being construed as the cause in causal explanations.

Finally, *causal relevance* may be read as a vague term in need of clarification before any conclusive determination can be made about mental properties. Hence, two consequences of this survey are that program explanations can be seen as a heterogeneous category with supervenient, mereological, or reductive relations as possible relations holding between program and process explanatory properties. But more importantly, the essence of program explanations is not in principle something Kim had not already considered as a possible relation between two supposedly complete causal explanations of a single event.

Conclusion

Jackson and Pettit argue that even if mental properties are causally inefficacious, they can be causally relevant by providing information that rival explanations citing efficacious properties are unable to convey. When the examples given by Jackson and Pettit to show the relevance of program properties are examined with explanatory exclusion in mind, they can be seen to be consistent with this principle. As it states that to accept both explanations all that is required is that there is “an appropriate account of how they are related to each other.”³⁶ If either supervenience or the mereological relation can account for the relevance of program explanations, then we have an acceptable relation in Kim’s terms. If program explanations do provide us with an account not considered by Kim, then that distinctness still has to be shown.

There are three general conclusions that can be drawn from the survey carried out in this essay. First, program explanations fail to refute the

³⁵ Donald Davidson, “Mental Events,” in *Essays on Actions and Events* (USA: Oxford 1980).

³⁶ Kim, “Mechanism, Purpose, and Explanatory Exclusion,” 257.

plausibility of the explanatory exclusion principle and rather, are consistent with the principle by being comfortably accommodated for within Kim's wider framework. Second, the issues raised by the debate surrounding explanatory exclusion and the causal relevance of higher-order properties can be plausibly seen to have influenced the direction of recent attempts to substantiate not only the causal relevance of such properties, including mental properties, but also the causal autonomy of higher-order properties. Finally, the examples proposed by Jackson and Pettit are heterogeneous with respect to the relations they exemplify and when considered as analogous with mental properties, do not vindicate the causal relevance of the mental one way or another. More accurately, any causal relevance attributable to a mental property will depend on the precise relation that exists between that mental property and its correlated physical property.*

Department of Philosophy, Macquarie University, Australia

References

- Block, Ned, "Do Causal Powers Drain Away?," in *Philosophy of Phenomenological Research*, 67 (2003).
- Davidson, Donald, "Mental Events," in *Essays on Actions and Events* (USA: Oxford 1980).
- Jackson, Frank, and Philip Pettit, "Functionalism and Broad Content," in *Mind*, 97 (1988), 381-400.
- _____, "Program Explanation: A General Perspective," in *Analysis*, 50 (1990), 107-117.
- _____, "Causation in the Philosophy of Mind," in *Philosophy and Phenomenological Research*, 50 (1990), 195-214.
- _____, "In Defence of Explanatory Ecumenism," in *Economics and Philosophy*, 8 (1992), 1-21.
- Kim, Jaegwon, "Mechanism, Purpose, and Explanatory Exclusion," in *Supervenience and Mind* (Cambridge University Press, 1993), 237-264.
- _____, *Supervenience and Mind* (Cambridge: Cambridge University Press, 1993).
- _____, *Mind in a Physical World* (London: Massachusetts Institute of Technology, 1998).
- _____, *Physicalism or Something Near Enough* (Princeton: Princeton University Press, 2005).
- Menzies, Peter, "The Exclusion Problem, The Determination Relation, and Contrastive Causation," in *Being Reduced – New Essays on Reductive Explanation and Special Science Causation*, ed. by Jakob Hohwy and Jesper Kallestrup (Oxford: Oxford University Press, forthcoming).

* I would like to thank Lise Marie Andersen for helpful comments on a draft of this paper as well as an anonymous referee for this journal for focusing my attention on issues I had overlooked while writing it. Also, thanks to Paolo Bolanos for encouraging my efforts.

- Pearl, Judea, *Causality: Models, Reasoning, and Inference* (New York: Cambridge University Press, 2000).
- Shoemaker, Sydney, *Physical Realization* (New York: Oxford University Press, 2007).
- Woodward, James, *Making Things Happen: A Theory of Causal Explanation* (New York: Oxford University Press, 2003).