

CDD: 128.2

CAUSAL EXCLUSION WITHOUT EXPLANATORY EXCLUSION¹

ANDRÉ FUHRMANN

Centro de Pesquisa
Universidade São Judas Tadeu
Rua Taquari, 546
03166-000 SÃO PAULO, SP
BRASIL

prof.afuhrmann@usjt.br

Abstract: The causal/explanatory exclusion argument is one of the principal weapons against the possibility of mental causes/explanations having genuine causal/explanatory power. I argue that the causal and the explanatory versions of the exclusion argument should be distinguished. There are really two arguments, one of them perhaps successful, the other one not.

Keywords: mind; causation; explanation; exclusion; physicalism; supervenience.

¹ This paper was written while holding a Heisenberg-Fellowship of the German Research Council (DFG) and while visiting the Universidade Federal do Rio de Janeiro with grants from CAPES and DAAD. Many thanks to my host Wilson Mendonça in Rio with whom I went through various versions of the paper. The paper was first read at the conference “Mental Causation/Causação Mental” in João Pessoa, Paraíba, in March 2002. Special thanks are due to the organizer of the conference, André Leclerc (UFPB), and to my commentator, Hilan Bensusan (UnB).

1. INTRODUCTION: "CAUSAL/EXPLANATORY" EXCLUSION

There are two versions of the exclusion argument and these versions are not always distinguished. Some philosophers explicitly hold, that the two versions are but two ways of presenting one argument.

Here are two quotes from the work of Jaegwon Kim. They summarize what Kim takes to be one, if not the most instructive, puzzle concerning mental causation.

This is "the problem of causal/explanatory exclusion": For any single event, there can be no more than a single sufficient cause, or causal explanation, unless it is a case of causal overdetermination (Kim (1996), p. 150.).

The general point I want to stress is this: the presence of two causal stories, each claiming to offer a full causal account of a given event, creates an unstable situation requiring us to find an account of how the two purported causes are related to each other. This is the problem of "causal/explanatory exclusion" (Kim (1998), p. 65.).

Why are there in each episode of mental-to-physical causation *two prima facie* distinct causal stories to be told? Suppose that some physical effect P is explained by way of a mental cause M . For example, my desire for fresh air triggers a causal chain which eventually explains the fact that a window turns in its hinges and ends up in a new position. Since P is a physical effect it is eminently plausible to assume that it has a causally fully sufficient physical history. Tracing back the causal history of P , we shall never need to introduce non-physical causes to explain why P came about. Thus, if Q is any such physical cause fully sufficient for bringing about P , we have in fact two answers to the question why P occurred. The one is that, by hypothesis, P occurred because of M , the other is that P occurred because of Q . This immediately raises the question as to how the two causes M and Q are related to each other. As it turns out, the space of possible answers is strictly limited.

First possibility: M and Q are only partial causes, if taken separately; they explain P more fully, if taken together. In that case neither M nor Q can be taken to offer a full causal explanation of P . This possibly does not apply in the present case since we assumed both M and Q to be both fully sufficient for P .

Second possibility: Both M and Q fully explain P . But M does so by way of explaining Q (or *vice versa*). In this case M and Q are arranged in a sequence. Example: Why did Peter lose the game? Because Pedrita moved the knight to B7. Why did Pedrita move the knight to B7? Because Peter had incautiously opened his right flank. So Peter lost the game because he had opened his right flank. Here we have a case of more than one full causal explanation for a given effect; but the two explanations are not rivals but can be chained to obtain a more detailed explanation. Again, this possibility may safely be discarded. The simplest reason is that there is nothing in the story told above that coerces us to assume that in each case of mental causation the mental cause M must not be simultaneous with the postulated physical cause Q . But if M and Q are simultaneous causes, then the one cannot cause the other.

Third possibility: We are presented with a genuine case of causal/explanatory overdetermination. Both M and Q are fully sufficient and independent causes of P . Such cases may occur but they are usually held to be extremely unlikely coincidences. It seems implausible to suppose that each case of mental causation involves such a coincidence of two separate and each on its own fully sufficient causes.

Fourth possibility: The classification of causes into disjoint classes, mental and physical, is bogus. On closer inspection M turns out to be a physical item which fits perfectly well into the physical causal chain of which Q is apart. So M may be just Q or some other physical cause arranged sequentially with Q . This is a live possibility only if we repudiate the idea of a distinct class of mental causes. If this possibility is real, mental causation is just physical causation.

Fifth possibility: M and Q are distinct but stand in some non-causal relation to each other and this fact dispels the tension between M and Q as two complete causal explanations for a single effect. The problem with this possibility is that it is difficult to plausibly construe the requisite relation between the two causes without inviting the objection that the causal line from M to P would then have to be viewed as parasitic on the process that leads from Q to P (or vice versa). But then there is really only one causal process involved here and the assumption of a genuine case of mental-to-physical causation proves ill-founded after all.

This is all familiar ground by now and sets the agenda for much of contemporary philosophy of mind. Identity theorists, like Kim, recommend that the observed tension created by the assumption of mental-to-physical causation should be turned into a *reductio* argument in favour of the identification of mental with physical causes. This *reductio* argument naturally proceeds by discounting the four alternative possibilities.

Much, of course, depends on how the tension is presented. In the presentation above I have taken much care to blur the line between causation and explanation. In a moment I shall try to pull these two aspects apart, distinguishing one version of the argument that draws on a principle of *causal* exclusion from another that relies on a principle of *explanatory* exclusion. Kim, I believe, is well aware that these principles are distinct. But he also believes that the two are conceptually linked in such a way that fundamentally the same tension will arise with either of them. It is for this reason that Kim speaks of the “causal/explanatory exclusion argument”. I shall critically examine the purported link in Section 5 below.

Explanationists, such as Lynne Rudder Baker and Tyler Burge, also assume that there is such a link between causation and explanation. For them, however, the link works the other way round. For them causation is a derivative of a certain type of explanation. Since they reject

the problem of explanatory exclusion they conclude that there can likewise be no problem of causal exclusion. For reasons that need not be explained here I feel little inclination to the view that explanation needs no metaphysical grounding. Instead, I shall argue that even if we take explanatory claims at face-value, thus interpreting them as asserting something about the “real” causal structure of the world, the issue of mental explanation is importantly distinct from the issue of mental causation. The view that there are genuine mental explanations of a causal type does not trigger a problem of explanatory or causal overdetermination.

2. TWO EXCLUSION ARGUMENTS

Let us begin by first stating the exclusion problem in purely causal terms, that is, without reference to explanation. I shall present the problem by showing the (likely) inconsistency of four assumptions.

(a) Suppose that a mental cause c causes a physical effect e . This is the assumption of *mental causation*. (b) Assuming that the physical world is *causally closed*, we know that there is also a physical cause c^* which is fully sufficient for e . (c) That physical cause c^* need neither be interpolated between c and e nor need it be a cause of c . That is to say, we may assume that c and c^* are *not sequenced*. As remarked above, the physical cause c^* may be simultaneous with the mental cause c – which is in fact the story typically told in supervenience theories of mental causation. (d) We assume further that mental and physical causes are *distinct*. (e) Finally, the assumption of *exclusion*:

if c causes e and c^* causes e , then either

- (i) c is the same cause as c^* , or
- (ii) c and c^* stand in a causal sequence leading to e , or
- (iii) c and c^* are only jointly sufficient to cause e , or
- (iv) c depends on c^* (or *vice versa*) in order to cause e .

But the assumptions of (a) mental causation, (b) causal closure, of (d) distinctness and of (c) no-sequencing, rule out each of the alternatives mentioned in the consequence of the exclusion principle:

- (i) is ruled out by the *distinctness* of mental and physical causes;
- (ii) is ruled out by the assumption that c and c^* are *not sequenced*;
- (iii) is ruled out by *closure*, since c^* is assumed to be a fully sufficient cause of e ; and finally
- (iv) is likely to make c a second rate cause of e , thus offending against the assumption that genuine *mental causation* occurs.

Thus the four assumptions (a-d) cannot be conjoined to the exclusion principle; at least one has to be given up.

Next let us try to transpose this argument to a contradiction into the language of explanation. Instead of cause and effect we shall now talk of explanans and explanandum, and explanation will replace causation. The attempted transposition – which I do not wish to endorse for reasons to be detailed in a moment – would be as follows:

The assumption of mental causation turns into the assumption of *mental explanation*: we suppose that a physical explanandum b is explained by some mental explanans a . We assume further a principle of *explanatory closure* of the physical domain: given that b is a physical explanandum, we should postulate that there is in principle some complete physical explanans a^* for b . It seems to me that the credentials of this principle are no worse than those of the corresponding causal closure principle. Whether or not a good metaphysical grounding is available, the principle of explanatory closure has at least the status of a deeply entrenched methodological maxim and it would be unfortunate if the practice of mental explanation required giving up this maxim. The scenario described so far does not require that the postulated physical explanans a^* must explain a or *vice versa*. So we have the required assumption of *no-sequencing* in the case of explanation as well. And we also assume that mental and physical

explanantia are *distinct*. Finally we formulate the required principle of *explanatory exclusion*.

if *a* explains *b* and *a** explains *b*, then either

- (i) *a* is the same explanans as *a**, or
- (ii) *a* explains *a** or *a** explains *a*, or
- (iii) *a* and *a** are only jointly sufficient to explain *b*, or
- (iv) *a* depends on *a** (or *vice versa*) in order to explain *e*.

But, as before, each of these alternatives is ruled out by the assumptions just made. Alternative (i) is ruled out by the hypothesis that *a* and *a** are distinct explanantia; the one makes for an explanation of *b* in mental terms, the other explains *b* in physical terms. Alternative (ii) may be ruled out by simply restricting attention to causal explanation. In this case *a* and *a** may refer to simultaneous occurrences between which no causal explanation can obtain. For example, the two occurrences may stand in a relation of supervenience to each other. But supervenience is not an explanatory relation but a relation that is itself in need of explanation. Alternative (iii) cannot obtain because we assume that *a** is a complete physical explanans, as postulated by the explanatory closure of physical explananda.

Alternative (iv) requires perhaps a little bit more comment. A likely candidate for the requisite relation of dependency between explanantia is some logical relation (in a wide sense). If so, then mental *a* would be logically derivative of physical *a**. But this is just to say that mental explanations do not constitute a *sui generis* domain of explanations; they are a mere corollary to physical explanations. I take it that this is more reduction than a believer in the autonomy of mental explanation is prepared to accept. So (iv) is likely to come into conflict with a sufficiently robust understanding of the thesis that there are successful instances of mental-to-physical explanation.

In any case, whatever explication of the notion of dependency may be offered, a defender of mental explanation must take care to construe it in such a way that mental explanation remains sufficiently independent from physical explanation. Opponents of genuine mental explanation typically argue that this challenge cannot be met. Thus, pending a better understanding of dependency in alternative (iv), there seems good reason to believe that the principle of explanatory exclusion as stated above is in conflict with the assumptions usually made when considering cases of mental-to-physical explanation.

The analogy between the causal and the explanatory exclusion problem is certainly tempting. However, I believe, as indicated above,

that it is also superficial because it trades on a seriously truncated view of explanation, completely ignoring its essential intensional and pragmatic aspects². Once these aspects are taken into view the explanatory version of the exclusion problem takes on the air of a philosophical make-believe.

3. THE FAILURE OF EXPLANATORY EXCLUSION

There is no problem of explanatory exclusion. The context relativity of full explanations leaves room for the massive occurrence of a plurality of explanations for a given explanandum.

A body is found in Crackenthorpe's garden³. As this is a rather uncommon phenomenon, it calls for an explanation, that is to say, an answer to the question as to why the body was found in Crackenthorpe's garden. At the beginning of the enquiry the answer "Because the railway passes nearby" is a perfectly good explanation. It integrates the further fact that only a few days before, a murder was witnessed in a passing train travelling along the track adjacent to the garden. Now the investigation proceeds on the assumption that the body was dropped from a passing train and thus the fact that the railways passes nearby the garden, though still true, will no longer be accepted as an explanatory fact. Instead, an explanation of the fact that a body was found in Crackenthorpe's garden will now be expected to contain some information as to why the victim was murdered and why it was dropped at this particular place along the track, and so on.

² I am not accusing Kim of yielding to the temptation of a superficial analogy. Kim has a well-considered view about the relation between (causal) explanation and causation and it is this view which convinces him that there is essentially only one exclusion problem which may both be cast in terms of explanation as well as in terms of causation. I shall turn to Kim's view in this respect in Section 4 below.

³ Agatha Christie, *4:50 from Paddington*, 1957.

For every explanation offered we can always go on asking further why-questions. When the murderer is finally revealed we may still want to know why he committed the crime. When we are told that he did it for greed, we may *backtrack* further by asking why he was such a greedy character. Or we may *sidetrack* by asking why he chose to strangle rather than to shoot the victim. None of these and many other questions receive an answer in Agatha Christie's novel. But this is not to say that the Crackenthorpe case eventually remains unresolved. The possibility of backtracking and sidetracking in response to any explanation offered does not show that no explanation is ever fully adequate. It only shows that the adequacy of an explanation is always relative to a given set of background beliefs and expectations. Given the right circumstances an explanation can provide a fully satisfactory answer to a why-question without needing to anticipate all relevant backtracking or sidetracking moves. The adequacy of an explanation is simply dependent on what the addressee believes and expects. This, I believe, is a commonplace fact about explanation. Controversy only arises over how the commonplace is to be spelt out.

Without entering into such controversy we can observe that different explanations of a given fact need not give rise to an unstable epistemic situation. Why did the baby cry? Here are two possible explanations: (i) because the water in the tub was too hot; (ii) because the mean kinetic energy of the H₂O-molecules in the tub had passed a certain threshold value. Although the temperature of water just is a function of its mean kinetic energy, this does not mean that the two explanations are always interchangeable, let alone that the first explanation is in any sense second rate and that only the second is genuine. In fact, for those who do not know of the relation between heat and molecular movement, that is to say, for most people, (ii) would not even pass as an explanation at all. Their beliefs are not such that (ii) could be received as an explanatory fact. And even those who do know of the microphysical basis of heat are likely to think of (ii) as a strange

and presumptuous way of indicating (1) as the explanation rightly expected under the circumstances.

If we try to do justice to this intensional aspect of explanation, we can quickly observe that it becomes difficult, if not impossible, to generate the exclusion problem for mental explanation.

Suppose we are confronted with a *mental explanation* of some physical explanandum *b*: a mental explanans *a* explains successfully in context *C* why *b* has occurred, or,

$$(1) \quad a \rightarrow_c b,$$

as I shall abbreviate in the sequel.

We have adopted above a closure principle to the effect that every physical fact has a complete physical explanation. But now we have to exercise some care as to how the necessary reference to a context of explanation is to be brought into play. Here are two possible ways of formulating the required closure principle:

Weak closure

For every physical explanandum *b* there exists a physical explanans *a** such that *a** explains *b* in *some* context *D*.

Strong closure

For every physical explanandum *b* there exists a physical explanans *a** such that *a** explains *b* in *every* context *D*.

Let us then try to create the aimed at tension, first by using the weak closure principle. Given the assumption (1), it follows by weak closure that the physical effect *b* has a purely physical explanans *a** in some suitable context *D*, i.e.

$$(2) \quad a^* \rightarrow_D b.$$

But the prongs of the fork thus constituted by (1) and (2) do not have the least tendency to exclude each other. As just illustrated, it is a ubiquitous phenomenon that a given fact is explained differently in different contexts. There is no danger of explanatory overdetermination here. If the realm of physical facts is only weakly closed under explanation, then no exclusion problem arises.

The situation is different, if we assume the strong closure principle. For in this case we may infer from (1) that there exists a physical explanans a^* which explains b in the same context in which the mental explanans a explains b , i.e.

$$(3) \quad a^* \rightarrow_C b.$$

Now we have indeed one explanation too many and the question naturally arises as to how the two explanantia a and a^* are related.

The simplest solution would be to render a and a^* identical. But on the plausible assumption that mental a and physical a^* cannot replace each other in all contexts of explanation, this possibility clearly does not exist. Similarly we can safely rule out the suggestions that a explains b only by way of explaining a^* or that a^* only explains b by way of explaining a , or that a and a^* successfully explain b only if taken jointly. Finally, what about the suggestion that a and a^* stand in some other dependency relation to each other? Suppose, as seems more plausible, that the mental fact a , cited as an explanans of b , depends on the physical fact a^* which also explains successfully b . Must we then not say that the success of the mental explanation $a \rightarrow_C b$ is only parasitic on that of the physical explanation $a^* \rightarrow_C b$, that only the latter offers the *real* explanation of why b occurred, whereas the former deserves the title “explanation” only by way of vaguely indicating the availability of a real

explanation? If this danger is real, as Kim and others have persuasively argued, then it undercuts our initial assumption that there exist genuine cases of mental-to-physical explanations. Thus, given strong closure, the exclusion problem may arise for mental explanation as it arises for mental causation.

4. STRONG EXPLANATORY CLOSURE, ACCESSIBILITY AND MERGING CONTEXTS

A problem of explanatory exclusion can be generated from the analogous problem of causal exclusion by assuming a strong version of the principle of explanatory closure of the physical domain. But this version of the principle is not very plausible.

How plausible is the strong closure principle that for every physical explanandum there exists a purely physical explanans which is fully satisfactory in every context of explanation?

As a preliminary observation, let us note that strong closure is not needed to subserve the thesis that physical facts occupy a special position in the family of facts. Weak closure suffices to assign to physical facts that special rôle that we have in mind when we affirm that, say, unlike biology, we never have to leave the realm of physics when we wish to explain why some physical fact obtains (provided that the fact in question admits of an explanation at all): relative to some appropriate context, there will always be some good physical explanation available. Although weak closure thus suffices to distinguish the family of physical facts, strong closure would of course even strengthen the distinction of the physical. But is such strengthening plausible?

Moving from weak to strong closure amounts in effect to embracing the thesis that for every physical fact to be explained there is a further physical fact which makes for a fully satisfactory explanation irrespective of the context of explanation. According to this thesis then, at least some explanations, in particular certain physical explanations –

explanations from physical explanantia to physical explananda – are not context-sensitive in the sense explained above.

Some such view seems to underly Davidson's argument for anomalous monism when he tries to draw a distinction between psychological and physical regularities. The former, he maintains, are at best rules of thumb, default laws, whereas the latter are "strict and without exception". Thus Davidson writes:

[P]sychological explanations are never full and sufficient; like most explanations, they are interest-sensitive, and simply assume that a vast number of (unspecified and unspecifiable) factors that might have intervened between cause and effect did not ((1995), p. 16).

Presumably we are invited here to draw the conclusion that physical explanations can be "full and sufficient" and that they are not defeasible in the way explanations in psychology typically are. However, in the same article Davidson treats the idea of a "full" physical explanation with a good portion of reservation (always putting it in scare quotes). Reporting Kim's view, he writes

The idea is that if physics does provide such 'full, sufficient' explanations, there is no room for mental explanations unless these can be (fully, strictly?) reduced to physical explanations. What can this strange principle mean? If we consider an event that is a 'full, sufficient' cause of another event, it must, as Mill pointed out long ago, include everything in the universe preceding the effect that has a causal bearing on it, some cross section of the entire preceding light-cone; and even then, if we take 'sufficient' seriously, we must assume perfect determinism ((1995), pp. 15-16.).

Note that Davidson is moving here from scepticism about the epistemological notion of a full explanation to scepticism about the metaphysical notion of a full cause. Of course, if the metaphysical notion of a full cause is suspect, then the corresponding epistemological notion of a full causal explanation cannot fare any better (though not *vice versa*). But there is no need here to rely on metaphysical suspicions. It suffices to remind ourselves of the fact that explanation is an essentially

epistemic enterprise and that facts that cannot humanly be entertained can play no rôle as either explanantia or explananda.

A fact cited as an explanans can only play that rôle given plausible background expectations to the effect that the world cooperates. It is true though that we can choose to increase the generality of an explanation – that is, to decrease its context dependence – by making contextually implied information explicit. But it seems implausible in the extreme to suppose that in giving explanations the reliance on context can be completely eliminated in this way – even if we restrict attention to explanantia in basic physical terms, and even if we waive the natural limitations of real explanation-givers and receivers.

The best reply, then, to proponents of strong closure is to issue a challenge to explain in what sense an explanation can be both context-independent and fully satisfactory – leaving no room, that is, for either back or side-tracking.

There are two strategies for creating a problem of explanatory exclusion which seem more promising than assuming the excessively strong closure principle without further ado. I shall only briefly and rather abstractly sketch these strategies without elaborating the details eventually required.

The first strategy assumes that contexts of explanations are not always independent. A good explanation in one context may continue to be good in another one. If so, then explanations given in one context may carry over – be “accessible” – in another context. This, I believe, is a familiar phenomenon. The question is, whether it can be pressed into service to generate the exclusion problem.

To see how this might be done, suppose again that mental *a* explains in context *C* physical *b*. Assuming weak closure we postulate a physical explanans *a** for *b* in some suitable context *D*. Now, in order to progress beyond this point we need to argue that the explanation in

context D carries over to context C , or *vice versa*. Let us pursue only the perhaps more plausible former possibility. Thus we assume that context D is accessible from C and so the explanation given in D remains satisfactory in C . In that case we would have two explanantia, a and a^* in the single context C for the single explanandum b ; whence the exclusion argument could take its course.

The problem with this strategy is that it is dangerously close to assuming the strong closure principle. For in order to generate a general exclusion problem for mental explanation, we need to assume in general that the physical explanations postulated by closure carry over to any context in which a mental explanation has been given. But weak closure together with this assumption entails a thesis which, for the cases of interest here, is just as strong as strong closure. The critical question to ask then is, why contexts in which physical explanations hold should be accessible to all other contexts, in particular to all those contexts in which successful mental explanations can be provided.

The second strategy employs the idea that contexts may be merged. This, again, seems a natural idea. If, for example, contexts are represented by sets of possible worlds, then their merging may be represented by intersection. More involved representations require more involved merging operations. But in principle, the idea of somehow composing contexts of explanation poses no conceptual difficulty.

Thus again, we may use weak closure in order to infer from $a \rightarrow_C b$ that $a^* \rightarrow_D b$. Now, if we could argue that explanations are preserved under merging contexts, then an exclusion problem would arise. For then we would have to face the situation that both a and a^* explain b in context $C + D$, the merging of C and D .

However, the intensional nature of explanation makes it highly dubious whether there exists any suitable notion of merging contexts such that explanations are preserved under merging. That is to say, in general there will be no guarantee that if $a \rightarrow_C b$, then, for arbitrary

contexts X , $a \rightarrow_{C+X} b$. So how could there be any such guarantee in the case at hand?

The most plausible version of the strategy will likely turn out to collapse into the first strategy of accessible contexts. The idea would be to argue that since the context D of the physical explanation is more “specific” than the context C of the mental explanation, the merging of $C + D$ carries the same information as D , i.e. $D = C + D$. That would provide a plain sense in which C is accessible from D and so the explanation given in D would carry over to context C . Then we would have again two distinct explanantia of a single explanandum in the context C thus engendering the threat of overdetermination. The difficulty with this strategy is but a simple variation of the question raised with respect to the first strategy: one would have to explain in which sense the information carried by a context of mental explanation can add nothing to the information carried by certain contexts of physical explanation such that the physical explanations must also be acceptable in contexts of mental explanation.

To reemphasize, I have discussed these strategies only very abstractly, pointing out where more detailed elaboration is necessary and which likely problems need to be addressed. The discussion is, of course, far from conclusive and I do not mean to preclude that the one or the other version of these strategies may not be developed to considerable plausibility. For now, however, I shall proceed on the working hypothesis that no convincing explanatory exclusion problem can be generated in this way.

5. EXPLANATORY REALISM AND CAUSAL EXCLUSION

Neither does the principle of explanatory realism help to proceed from causal to explanatory exclusion.

For Kim, the problem of explanatory exclusion does not derive by analogy from the problem of causal exclusion but because of a certain

view he takes about the metaphysical grounding of causal explanations. To be successful causal explanations must be externally grounded in causal relations. This is Kim's answer to what he takes to be one of the two fundamental questions for theories of explanation:

The Metaphysical Question: When G is an explanans for E , in virtue of what relation between g and e , the events represented by G and E respectively, is G an explanation for E ? What is the objective relation connecting events, g and e , that grounds the explanatory relation between their descriptions, G and E ? (Kim, (1994), p. 56.)

One might deny that any such objective and external relation need to exist for an explanation to be successful. But if we take causal explanations at face value – and why should we not? – then they do seem to claim the existence of certain causal facts external to what agents believe and expect. In other words, causal explanations seem committed to at least this much

Causal Realism:

If a explains b (in some appropriate context), then there are events e_a and e_b , associated with the explanans a and the explanandum b respectively, such that e_a causes e_b .

For example, if I explain that Wilson is now rich because he won at the bingo, then I certainly claim that certain events have occurred – a significant increase of the deposits in his bank account, for example – and that these events have been caused – perhaps in conjunction with certain other events that I can presume to be known – by something that had occurred on a certain day in a certain room where someone suddenly shouted “bingo!”. Rejecting causal realism in this sense could only be taken as signalling commitment to an error theory or a fictionalist approach to causal explanation – options which I shall not discuss here.

Now, how is the fact of mental explanation supposed to create a problem? Kim writes that the explanatory exclusion problem “arises from the very notion of a causal explanation and what strikes me as a perfectly intuitive and ordinary understanding of the causal relation” (Kim (1998), p. 67).

For Kim, the very notion of a causal explanation entails the minimally realist view just mentioned; and the ordinary understanding of the causal relation rules out massive causal overdetermination. As I shall argue now, we can follow Kim in this and yet refuse to acknowledge any problem arising out of the combination of these views.

Suppose that we encounter two explanations in distinct contexts of the same fact b from distinct explanantia a and a^* ,

(1) $a \rightarrow_c b$, and

(2) $a^* \rightarrow_D b$.

According to the thesis of causal realism, these explanations must be grounded in causal facts,

(3) $c_a \rightarrow e_b$, and

(4) $c_{a^*} \rightarrow e_b$.

(Where c_a stands for the causing event associated with the explanans a , and \rightarrow represents causation, which is, of course, not relative to epistemically determined contexts.)

This is, of course, the familiar situation which satisfies the antecedent of the causal exclusion principle. It is for this reason that Kim concludes that “metaphysics won’t go away” and that the causal exclusion problem arises from premisses like (1) and (2).

What is certainly true is that given the magic triangle constituted by (3) and (4) we owe some account of the relation between the events c_a and c_{a^*} , and that the consequent of the causal exclusion principle

presumably lists all available possibilities. But are these possibilities ruled out by the assumptions made, as needs to be the case in order to create a tension between the two explanantia a and a^* ?

Here are, again, the four possibilities: (i) $c_a = c_{a^*}$, (ii) c_a causes e_b only by way of causing c_{a^*} first, or c_{a^*} causes e_b only by way of causing c_a first, (iii) only jointly do c_a and c_{a^*} suffice to cause e_b , (iv) c_a depends in some non-causal way on c_{a^*} or vice versa. Possibility (ii) may be safely discarded. It is not credible to suppose that a situation of the kind described in (1) and (2) can only obtain, if the causes underlying the explanantia are chained. I shall simply pass over (iv) here – for the sake of the argument it may be assumed that this too is not a viable possibility. On possibility (iii) I shall comment in a moment.

Although we have assumed, with good reason, that the explanantia a and a^* must be rendered distinct, there is not the slightest reason to suppose that therefore the causes c_a and c_{a^*} , which ground a and a^* , must also be distinct, thus ruling out possibility (i). Assuming that they are, amounts in effect to assuming that there exists a bijection between the domain of explanation and the domain of causation, that explanantia and explananda just are causes and effects, or at least stand in a one-to-one relation to each other. This is an assumption much stronger than causal realism and there is little to commend it. Thus, pending a decision on possibility (iii), we may simply maintain that both a and a^* successfully explain b in their respective appropriate settings, that they do so in virtue of an external causal relation, as required by causal realism, and that the appearance of causal overdetermination is best resolved by assuming that a and a^* are grounded in the same single event that causes the event referred to by the explanandum b .

To revert to our earlier example, the fact that the temperature in the tub is well above the baby's body temperature is a distinct explanans from the fact the water molecules in the tub move with a certain velocity. But the distinctness of these facts *qua* explanantia does not preclude that they are both grounded in the same event. In fact our ordinary

understanding of causation, as given expression in the causal exclusion principle, together with the assumptions that are reasonable to make will counsel us to believe in distinct causal events only in exceptional circumstances, if ever. But if the causes underlying the distinct causal explanations (1) and (2) are identical, then the two explanations need not exclude each other. To put it in a somewhat pointed way: The best reason for supposing that the causes c_a and c_a^* are identical is the causal exclusion argument as put forward by Kim and others. Precisely because the *causal* exclusion problem is real, the explanatory exclusion problem is bogus.

Let us finally discuss possibility (iii), that the events c_a and c_a^* only jointly suffice to cause e . I have sketched above a sense in which causal explanantia can be complete and sufficient in appropriate circumstances: in such circumstances they can provide a fully satisfactory answer to a given why-question. Everyone who understands a why-question raised in a certain context knows what kind of answer would lay the issue to rest in that context. In that sense I take the notion of a complete and sufficient causal fact (explanans) to be unproblematic.

Problems soon arise when we move from facts to events. What is the event that was the full and complete cause of a certain coconut falling off the tree? Perhaps it was the wind moving the tree at that moment. But had the coconut not yet reached a certain weight and had the wind blown from a slightly different angle, the coconut would not have fallen off the tree then but, perhaps, a few days later. Thus the complete cause that made the coconut fall is either an event that had the given effect only because it occurred in a world collaborating towards that effect or it is a very huge and complex event.

6. CONCLUSION

Explananda may well be overdetermined by the facts, provided the contexts of explanation are different. This does not preclude that the

events grounding causal explanations are governed by a principle of causal exclusion.

As a piece of philosophical make-believe the credibility of the problem of explanatory exclusion relies much on a certain way of putting it – on choosing a certain terminology which furthers the escalation of a problem about causation to a problem about explanation. In the circumstances a de-escalation might be called for to restore proper balance. What we need is a plausible way of describing the place of mental facts in causal explanations without compromising the basic physicalist conviction that creatures with a mental life are part of the physical world.

One way of obtaining such a description is by distinguishing between facts and events. As I used these words for the purpose at hand, facts are individuated by their rôles in explanations: they are potential explanantia and explananda. Events are the kind of entities that underlie facts: they are potential causes and effects. Given this distinction between facts and events, we are free to maintain the following:

- mental and physical facts are distinct;
- mental facts can causally explain physical facts in appropriate contexts;
- such mental-to-physical explanations need not be second rate and can be as complete as explanations in general can be;
- every physical fact has a full physical explanation in some context; and
- facts, whether mental or physical, are only exceptionally explanatorily overdetermined in a single context.

Moreover, causal explanations may be interpreted as making factual claims. That much causal realism does not give rise to an unstable situation as long as the contexts of mental and physical explanations are distinct. For if two distinct facts explain a single third fact in varying

contexts, causal realism counsels us to assume that the two causal explanations are being subserved by a single episode of causation where the single causing event underlies two distinct facts. The metaphysical question whether there is a genuine category of causally efficacious mental *events* is largely independent from the question as to whether mental facts can be causally efficacious in the sense of being capable of serving as explanantia in causal explanation.

REFERENCES

- DAVIDSON, D. (1995) "Thinking Causes", in J. Heil & A. Mele (eds.) *Mental Causation* (Oxford University Press, Oxford).
- KIM, J. (1994). "Explanatory Knowledge and Metaphysical Dependence", *Philosophical Issues*, 5.
- . (1996). *Philosophy of Mind* (Boulder, Westview Press).
- . (1998). *Mind in a Physical World* (Cambridge, Mass., MIT Press).