

7-2015

Mental Causation Is Not Just Downward Causation

Jeff Engelhardt
Dickinson College

Follow this and additional works at: http://scholar.dickinson.edu/faculty_publications

 Part of the [Philosophy Commons](#)

Recommended Citation

This article is brought to you for free and open access by Dickinson Scholar. It has been accepted for inclusion by an authorized administrator. For more information, please contact scholar@dickinson.edu.

Mental Causation Is Not Just Downward Causation

Abstract: According to a popular model of mental causation, an irreducible mental cause M1 brings about an irreducible mental effect M2 by bringing about M2's supervenience base (ground, realizer, etc.), P2. Call this "the Downward Causation View". This paper raises doubts about the Downward Causation View on grounds that M1 doesn't cause M2 immediately and there is no causal chain from M1 to M2. *Prima facie*, then, M1 doesn't cause M2 on this view. But a theory of mental causation ought to account for how some mental phenomena cause other mental phenomena; so rival theories are to be preferred. After setting out the problem, I consider replies; all fail.

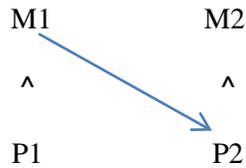
Keywords: mental causation, causal exclusion, downward causation, causal explanation, causal powers.

1. Introduction

We use causal explanations as defeasible indicators of responsibility. If Jill's anger caused her to push Jack down the hill, then she's responsible for it. If this practice is to be coherent, it must be that mental causes have physical effects *and* that mental causes bring about intermediary mental effects. Jill's anger caused a planning process; the planning mechanism produced her intention; her intention caused a motor process; and so on. Just as we need an account of mental-physical causation, we need an account of mental-mental causation.

Since mental causes and effects seem to depend on physical phenomena for their existence, many have thought it plausible that mental-mental causation proceeds 'through' physical phenomena. Thus, there is a popular model of mental causation according to which an irreducible mental cause M1 brings about an irreducible mental effect, M2, by bringing about something physical on which M2 non-causally depends—its supervenience base, physical ground, realizer, etc.—P2.¹

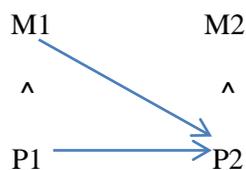
¹ See for instance Ned Block (2003: 135) "M can only cause M* by causing [M*'s 'vertical determiner'] P*"; Karen Bennett (2007: 326) "M₁ could cause M₂ by causing its supervenience base P₂ to be instantiated. It is independently plausible to think that this is the only way to cause a supervening property to be instantiated."; Ausonio Marras (2007: 310) "In order to cause M*, M must surely bring about the conditions upon which M* depends [i.e. P*]"; and, Sven Walter (2008: 676) "in order for M to cause M*, it must cause M*'s supervenience base P*".



The broad metaphysical position at issue here is non-reductive physicalism. It is non-reductivism because the mental phenomena in the model are irreducible. In this paper, I'll take it that a mental phenomenon is irreducible just in case it isn't identical to any physical phenomenon. The position is nonetheless physicalist provided each mental phenomenon depends for its existence on some physical phenomenon or phenomena. Here, M1 depends for its existence on P1 and M2 depends on P2. This dependence relation, whatever it is, must be 'intimate' or 'strong' enough that physical phenomena ground the existence and nature of mental phenomena.

Otherwise, the model makes no assumptions about the ontological nature of M1, P2, or M2 as causal relata. They may be events, tropes, properties, etc. M1 and M2 represent *whatever* the relata of causal relations are, and P2 represents M2's physical ground, supervenience base, etc. *whatever* it is. The model can—and has been—adopted by those with varying ideas about the causal relata, the mental-physical relation, etc. Call it "the Downward Causation View".

Since it is assumed that P2 also has some physical cause—call it P1—the Downward Causation View faces a concern about overdetermination. If causes fully determine their effects, it seems as though P2 is *fully* determined in the model *twice*.²



Thus, many philosophers have tried to explain how P2 can have two causes without accepting that the determination of P2 is akin to the overdetermination of a dictator's death from two separate shots in a firing squad. The unquestioned assumption in the literature seems to be that if M1 causes P2 and P2 is not problematically overdetermined, then there can be a non-reductive physicalist account of how M1 causes M2.³ This paper challenges this assumption

² See, for instance, Kim (1998), Block (2003), Bennett (2007), Wilson (2009, 2011).

³ Bennett (2003), MacDonald and MacDonald (2006), Wilson (2009), Ney (forthcoming).

by questioning the nature of the relation proposed: it is not a *causal* relation between M1 and M2.⁴ Rather, there is a causal relation between M1 and P2, and then there is a *non-causal* relation between P2 and M2. This shouldn't be regarded as a causal relation. Although some might accept that psychological explanations are not causal explanations (see, e.g. Kim 2007), unless one explicitly endorses an alternative, we should take the claims of those who favor the Downward Causation View at face value: M1 is supposed to *cause* M2 by causing P2. An advocate of downward causation may of course retreat to the claim that M1 doesn't *really* cause M2, but we should read this as a retreat away from her initial position. The Downward Causation View claims to preserve the appearance that some mental phenomena cause others; the problem discussed here suggests that in this the Downward Causation View fails. There is no recognition of this problem elsewhere in the literature.⁵

If the challenge raised here succeeds, the non-reductivist should aim to make sense of immediate mental-mental causation. She may appeal, for instance, to difference-making or causal proportionality to spell out how M1's influence on M2 needn't rely on an intermediary relation between M1 and P2. (See Menzies 2003 for a difference-making approach to mental causation; see Yablo 1992 for an appeal to causal proportionality.) For those with non-reductivist commitments, then, this paper speaks in favor of these alternatives; for reductivists, this paper challenges one of the more popular models of non-reductivist mental causation.

Although I'm not aiming to address worries about the overdetermination of physical effects like P2 in this paper, the conclusion argued for here has clear bearing on how we should think of that problem and its resolution. On the one hand, if there is no downward causation, then M1 doesn't in fact cause P2, and P2 won't be causally overdetermined. It has just the one cause, P1. On the other hand, if M1 doesn't cause M2 'through' its physical ground, P2, it

⁴ Note that if one denies that M1 causes M2 'through' P2, then concerns about the overdetermination of P2 by M1 and P1 may be misplaced. Moreover, considerations similar to those raised here raise questions about whether P1 causes M2. If it doesn't, then there should be no concern about whether M2 is overdetermined by M1 and P1. Without these overdetermination concerns, many formulations of the problem of causal exclusion are misguided. See Zhong 2011, 132-138 for discussion of these various formulations. Thanks to an anonymous referee for this journal for pointing this out to me.

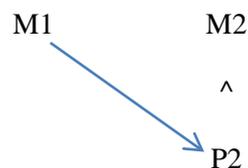
⁵ Note, for instance, that it is discussed in neither the *Stanford Encyclopedia of Philosophy* entry "Mental causation" (Robb and Heil 2013) nor the *Philosophy Compass* survey "Mental causation" (Bennett 2007).

remains unclear how to ensure that irreducible mental-mental causation is consistent with the physical determination of all mental phenomena.

2. The Problem with Downward Causation Views

The proponent of downward causation supposes that mental-mental causation involves the mental cause's bringing about the physical ground of the mental effect. But this runs contrary to an intuitive and popular idea about causation. It's commonly supposed in the causation literature that C causes E only if C causes E immediately or there is a causal chain leading from C to E.⁶ (See, e.g., Lewis 1973, 563; Mackie 1974, 46) If C doesn't cause E immediately and there is no chain of causes from C to E, then *prima facie* C is not a cause of E at all. And if C isn't a cause of E at all, then there is no causal relation between C and E. If C doesn't cause E, there may be some other relation between C and E, of course—if E is the statue and C is the bronze constituting it, C may *non-causally* determine E's occurrence and nature—but if the aim is to cite E's cause, C won't do.

But (for reasons discussed further below) proponents of the Downward Causation View do not suppose that P2 causes M2. M1 is a cause of P2 but P2 is not a cause of M2. M1's causing P2, then, is not a link in a causal chain leading to M2. There is, rather, no such chain.



We should ask, then: why think that M1 can cause M2 by causing P2? Our presumption should be against it. On the face of things, then, there is no downward causation. If so, then a theory that characterizes the relationship between M1 and M2 this way seems to fail in giving an account of how some mental phenomena cause others.

3. Replies

3.1 The chain of determination reply

⁶ Note that since this is given as a necessary condition on causation, those who reject the transitivity of causation may accept this condition and just deny that it's sufficient. (See Hall 200, Hitchcock 2001) Advocates of downward causation need to accept that causal chains suffice for causation while adding that partly non-causal chains of determination also suffice.

One obvious reply to this objection is to say that the relationship between P2 and M2 is sufficiently *like* a link in a causal chain that it's good enough to ground a causal relation. The reply calls for a revision in the necessary condition on causal relations given above. It's not that C causes E only if there is a causal chain from C to E, but only if there is *a chain from C to E with links sufficiently like causal links*. Causal chains 'still' ground causal relations, of course, since their links are sufficiently like themselves, but in addition, this revision permits the advocate of downward causation to claim that the non-causal dependence between mental and physical phenomena is sufficiently like a link in a causal chain. The conditions for being sufficiently like a causal link would of course have to be specified, but there's little reason to consider the possibilities here. For no matter the specification, this reply forces the non-reductive physicalist either to forfeit physicalism or to make *ad hoc* stipulations.

Why would it forfeit physicalism? If the relation between P2 and M2 is sufficiently like a causal link to ground causal relations, then it should ground a causal relation between P2 and M2. But the non-reductive physicalist cannot permit that P2 causes M2. By her commitment to physicalism, the non-reductivist needs each of her mental phenomena to depend on some physical phenomenon (or phenomena). But not just any dependence will do. To draw on a common metaphor, mental-physical dependence is supposed to guarantee that mental phenomena are 'nothing over and above' physical phenomena. However one understands the metaphor, though, causal dependence doesn't supply the needed guarantee. The swollen lip is indeed something 'over and above' the clumsy kiss that caused it.⁷ So the advocate of the Downward Causation View can't accept that P2 causes M2 on pain of forfeiting her physicalism.

One might hope to avoid this commitment by accepting that the relation between P2 and M2 doesn't ground a causal relation *on its own* but add that this doesn't prevent it from playing a 'supporting role' in the causal relation that allegedly holds between M1 and M2. This strategy asks us to distinguish between what it takes to 'get a causal relation started' (a

⁷ The possibility that the mental-physical dependence relation is a causal relation is hardly ever even entertained. It is not mentioned in the *Stanford Encyclopedia of Philosophy* entry "Physicalism" (Stoljar, 2009), in the *Philosophy Compass* survey "Can physicalism be non-reductive?" (Melnik, 2008), or in the *Philosophy Compass* survey "Defining physicalism" (Ney, 2008). Moreover, of those who do entertain this possibility, most reject it. See, e.g. Kim 1998, 47-50, Bennett 2003, 479. John Searle (Searle, 1992: 1) and E.J. Lowe (Lowe, 2008: 31, 48-57) are notable exceptions. It remains controversial, however, just how to interpret their suggestions and whether they're compatible with physicalist commitments.

causal relation) and what it takes to ‘keep a causal relation going’ (anything sufficiently like a causal relation). One then adds that the non-causal relation between P2 and M2 meets the latter standards but not the former. The necessary condition on causal relations is thus further revised: C causes E only if (i) C causes some D and (ii) there is a chain from D to E with links sufficiently like causal links. P2 doesn’t cause M2, then, because there’s no D of which P2 is a cause and from which M2 somehow results. M1 does cause M2, however, because (i) M1 causes P2 and (ii) there is a chain from P2 to M2 with links sufficiently like causal links. One might think this is the case when Chitra’s painting causes the canvas to be scarlet *and thereby red*. *Prima facie*, Chitra’s painting caused the canvas to be red. Why? Because (i) the event of Chitra’s painting causes the canvas to be scarlet, and (ii) since scarlet and red stand in the determinate-determinable relation, the canvas’s being scarlet determines its being red.

But while this view has some intuitive plausibility, when we look ‘beneath’ the loose, everyday conception, it is mysterious and *ad hoc*. The non-reductive physicalist must accept that there is *some* relation between P2 and M2 that doesn’t at all involve M1. This relation underwrites M2’s irreducible but nonetheless physical nature, as is required by the non-reductivist’s physicalism. It is this relation that is recruited to ‘complete’ the relation between M1 and M2. On the reply under consideration, this relation is non-causal. That is, this reply takes it that the relation between P2 and M2 (as between *being scarlet* and *being red*) is a non-causal relation. Thus, the reply says that there is a causal relation between M1 and P2 and a non-causal relation between P2 and M2, and yet *when we conjoin these relations, the resulting relation is just a causal relation*. The non-causal relation has somehow been ‘absorbed into’ or ‘recruited by’ the causal relation. One would have expected the resulting relation to be partly causal and partly non-causal, but the reply denies this. The whole in this case is alleged to be neither the sum of its parts nor more than their sum; it is less: the non-causal relation ‘disappears’. If this is what happens when causal and non-causal relations are conjoined, we should want to know what happens to the non-causal relation and why.

But one can’t really say that the non-causal relation *disappears*, of course. It’s needed to establish that M2 is physical. This reply, in other words, needs it to be that P2 both does and doesn’t non-causally determine M2. When M2’s physicality is in question, it needs P2 to non-causally determine M2. But when asserting that M1 causes M2, it needs that non-causal determination to ‘go away’. If this isn’t an outright contradiction, it’s a mystery. And if one

doesn't accept that the non-causal relation 'goes away', it's mysterious how a relation that's one part causal and one part non-causal may nonetheless be causal as a whole. Either way, the advocate of this reply owes us an explanation. But it seems the only answer the advocate of the Downward Causation View can offer us is that this maneuver preserves her theory. That is, this reply is *ad hoc*.

But an advocate of this reply, the chain of determination reply, might not find this satisfying. She may say in rejoinder that given the familiarity and ubiquity of cases like Chitra's painting, there's no need to look 'beneath' them for a better understanding of mental causation. On the contrary, such cases serve as models for mental causation, showing that at least some non-causal determination relations are so 'tight' that if a cause (Chitra's painting) brings about one relatum in a non-causal determination relation (the canvas's being scarlet), it thereby also stands in a causal relation with the other relatum of that non-causal relation (the canvas's being red).⁸ Far from 'disappearing', the non-causal determination relation here plays a crucial role. It's because the determinate-determinable relation between *being scarlet* and *being red* is so 'intimate' that any cause of one is also a cause of the other. The 'intimacy' of the non-causal determination relation explains the intuition that Chitra's painting scarlet is also painting red; similarly, *mutatis mutandis*, for M1's causing P2: given the intimacy of the mental-physical relation, it is also a causing of M2.⁹

I find these points attractive, but not as applied to the chain of determination reply. In the next section, I'll be considering a reply according to which the P2-M2 relation is very *unlike* a causal relation. I think the points given in the rejoinder are most plausible when taken in support of that view. Here, with the chain of determination reply, we are considering a different approach.

The claim here is that the M1-M2 causal relation is grounded in the M1-P2 causal relation and the P2-M2 non-causal relation, where the latter is alleged to be so similar to a causal relation that it can play the role of a causal link in the chain of determination from M1 to M2. This approach bears a considerable burden of proof, for it accepts that (i) the P2-M2 relation is not a causal relation and (ii) in general, non-causal relations are not also causal

⁸ See Bennett 2008, 287-292 for explicit appeal to the 'tightness' of non-causal determination relations. Many others appeal to this idea using other terms. See, e.g., Bennett 2003, Witmer 2003, Marras 2007.

⁹ Thanks to an anonymous referee for this journal for articulating this line of thinking very clearly and pushing me to develop an explicit reply.

relations (of course). As a result, it must ask us to reject the intuitive, traditional view that there is a causal relation between C and E just in case C causes E immediately or there is a chain of causal relations linking C and E. As grounds for this rejection, the chain of determination reply appeals to similarities between causal and non-causal relations.

Taking the rejoinder at hand in this context, it should tell us what makes the non-causal relation so similar to a causal relation. But its answer seems to do the opposite: the intimacy of non-causal relations *distinguishes* them from causal relations. Although effects depend on their causes, this relation is not as ‘tight’ or ‘close’ as that between a determinate like *being scarlet* and a determinable like *being red*. Where C causes E in actuality, there are possible worlds in which C occurs but E doesn’t (e.g. the clumsy kiss occurs but it doesn’t lead to a swollen lip); but scarlet can’t be instantiated without red. Instances of scarlet metaphysically necessitate instances of red, but causes don’t metaphysically necessitate their effects.

Indeed, advocates of the Downward Causation View have themselves highlighted this exact difference in defending their view against the charge of overdetermination. Where physical P1 non-causally determines mental M1 and M1 and P1 both causally determine P2 (as on page 1), why isn’t P2 overdetermined? Many, many advocates of the Downward Causation View claim that it’s because non-causal determination relations are so ‘tight’, ‘intimate’, etc., and many say explicitly that this distinguishes that relation from causation.¹⁰ For many think that if executioner E1’s shot causes executioner E2 to shoot and both shots hit the dictator at the same time, the dictator’s death is overdetermined. But if the particles that compose E1’s shot non-causally determine E1’s shot and both the particles and the shot hit the dictator at the same time, *this* doesn’t overdetermine his death. That is, many advocates of downward causation think that causal relations are not intimate or tight in the way that non-causal determination relations are, so that causal relations don’t ‘save’ an effect from being overdetermined but non-causal determination relations do. Where she appeals to the intimacy of non-causal relations, then, the advocate of the Downward Causation View doesn’t—by her own lights—give us grounds for thinking that causal and non-causal determination relations are similar. Rather, she appeals to a difference between the two. I take it, then, that this rejoinder doesn’t save the chain of determination reply: it offers no reason at all for thinking that non-causal relations are sufficiently similar to causal relations.

¹⁰ See, e.g., Bennett 2003 and 2008, Witmer 2003, Marras 2007.

3.2 The pre-philosophical conception reply

But as suggested above, considerations of ‘tightness’ are more helpful in the context of a different reply. One may say that the link between a physical effect and the mental effect it grounds is so ‘tight’ that causing the physical ground is enough to cause the mental effect. Like the reply above, this suggestion maintains that M1 causes P2 and then P2 non-causally determines M2; and, as above, the claim here requires a change to the standing conception of causation in the philosophical literature. But instead of saying that the P2-M2 relation is *like* causation, this reply emphasizes a way it is unlike causation: it is distinctively ‘tight’ or ‘intimate’. If the reply is to work, it must argue that the conception of causation in the literature is mistaken: it’s not true that C causes E only if C causes E immediately or there is a causal chain from C to E. The strategy here is to show that the philosophical conception is at odds with our pre-philosophical, ‘everyday’, or ‘intuitive’ conception of causation. It’s assumed that on the pre-philosophical conception, we would say, for instance, that Chitra caused the canvas to be red *by causing it to be scarlet*. Taking it that we should modify the philosophical conception to fit its pre-philosophical counterpart, we’re to accept that downward causation is genuine causation. M1, then, may cause M2 by causing P2. I call this “the pre-philosophical conception reply”.

If this reply is to serve as defense for the Downward Causation View, it requires a metaphysical specification and some argumentation. For we might agree that the philosophical conception should fit the pre-philosophical view but deny that the Downward Causation View provides the model that best fits the pre-philosophical conception. The standing philosophical conception might in fact be the best fit for the intuitive conception. In other words, the pre-philosophical intuitions on their own don’t establish the Downward Causation View; they leave much interpretation and modeling to be done. It’s *prima facie* dubious, after all, that a pre-philosophical conception of Chitra’s painting would include the metaphysical determinate-determinable relation between *being red* and *being scarlet*. If the pre-philosophical conception reply is to succeed, it should provide a metaphysical specification to link the intuitive conception of Chitra’s case to the Downward Causation View.

The proper subset account of the mental-physical relation seems well-suited for this purpose. On Jessica Wilson's subset account, every property is associated with certain causal powers, and the token powers associated with any given mental property are a subset of the token powers associated with the physical property on which that mental property non-causally depends.¹¹ (Wilson, 1999, 2009, 2011) Similarly, the token causal powers associated with *being red* are a proper subset of the powers associated with *being scarlet*. Thus, when Chitra's painting causes the canvas to be scarlet, it brings about a set of causal powers; since the powers associated with *being red* are a subset of those powers, Chitra has also caused the canvas to be red. As concerns mental-mental causation, we may say the same. When M1 causes P2, it causes the powers associated with P2 to be entokened. Since the powers associated with M2 are a proper subset of the powers associated with P2, M1's causing P2 suffices to cause M2 as well. Voila. In causing P2, M1 has caused M2 as well.

With the proper subset account offered as a metaphysical gloss, the pre-philosophical conception reply takes shape. The Chitra case has some intuitive force—I myself am inclined to say that Chitra's painting caused the canvas to be red. The proper subset view provides a plausible account of the case, and then one appeals back to our original intuitions about the case to motivate calling the M1-M2 relation in the model an instance of causation. Where the model depicts downward causation, we're impelled to accept that downward causation is in fact causation.

But there are at least two problems here. The first is that if the reply is thus posed, it doesn't shoulder its full burden. This kind of reply must show that we ought to reject the standing philosophical conception of causation on grounds that it is inconsistent with the pre-philosophical conception. But the inconsistency isn't obvious, and the pre-philosophical conception reply can't take it for granted. It must be shown. And that requires showing that there is no intuitively viable model of Chitra's case (or other intuitive cases) that is consistent with the standing philosophical conception of causation. All the reply has done so far is to claim that a specification consistent with downward causation is *one way* to model Chitra's case. We must be shown that there is *no way* to model Chitra's case without rejecting the philosophical conception of causation. Otherwise, the Downward Causation View has no

¹¹ One may find similar views in Shoemaker 2001, 2013 and Pereboom 2002.

'dialectical leverage' for denying that C causes E only if it does so immediately or there is a chain of causes from C to E.

Can Chitra's painting be modeled in a way that's consistent with this claim? Yes. First, one may say that the canvas's being scarlet is identical to the canvas's being red. Thus, Chitra causes the canvas to be red by causing it to be scarlet because causing it to be scarlet *just is* causing it to be red. In cases of mental-mental causation, the claim would be that mental phenomena are identical to physical phenomena. Thus, $P2 = M2$ and when $M1$ causes $P2$, it thereby causes $M2$. $M1$ on this view is an immediate cause of both $P2$ and $M2$, and thus this view is consistent with the claim that causes bring about effects either immediately or through a causal chain.

But the non-reductive physicalist can't accept this. For she denies that $P2 = M2$. She may still claim, however, that $M1$ causes both $P2$ and $M2$ immediately. As noted in the introduction above, several non-reductive views accept this. Indeed, the proper subset account offered above seems to motivate this claim. Recall that the pre-philosophical conception reply has two steps. First, it offers a case in which it's intuitive to say that some cause C brings about some effect $E2$ *by* bringing about some other effect $E1$ and in which there is a non-causal relation between $E1$ and $E2$. Second, it gives a metaphysical specification of that case that posits downward causation. It's the pressure to accept downward causation as a model of the case that is supposed to motivate the Downward Causation View and the rejection of the philosophical conception of causation. But if the metaphysical specification doesn't posit downward causation, there's no such pressure. And, arguably, the subset account doesn't posit downward causation. If not, the reply fails.

Call the powers associated with *being red* powers 1, 2, and 3; let the powers associated with *being scarlet* include these as well as power 4. When Chitra causes the canvas to be scarlet, she brings it about that powers 1, 2, 3, and 4 are all entokened. If we take it that her painting caused these entokenings immediately, it follows that her painting caused the entokenings of powers 1, 2, and 3 immediately as well. That is, her painting immediately caused the canvas to be red. It caused the canvas to be both red and scarlet immediately. And it doesn't matter much if her painting didn't bring about these effects immediately; the point is that *however* her painting caused the canvas to be scarlet, it caused the canvas to be red in the same way. The

causal relation between Chitra's painting and the canvas's being scarlet should be the same as that between her painting and the canvas's being red. If the former relation is consistent with the standing philosophical conception of causation, then so is the latter. Turning to mental-causation, we should say the same. If M1 immediately causes the tokens associated with P2, then since the tokens associated with M2 are a subset of these, M1 also causes these immediately. That is, M1 causes both P2 and M2 immediately.

The arguments given here don't say that the Downward Causation View should be rejected simply because it has competitors. We have to keep the reply's context in mind. It accepts that the Downward Causation View is inconsistent with the standing philosophical conception of causation and alleges that we must revise the latter in a way that accommodates downward causation. In an attempt to justify the accommodation, the reply offers (1) cases in which it's somewhat intuitive to say that a cause C brings about an effect E2 by bringing about some other effect E1 and (2) a specification of the case in which there is (i) a causal relation between C and E1 and (ii) a non-causal relation between E1 and E2. Together, 1 and 2 suggest that the pre-philosophical conception of causation accommodates downward causation; the reply then charges that the philosophical conception should too. But we have no reason to make the accommodation if the cases on offer in 1 are easily shown to be consistent with the standing philosophical conception. And, as we've just seen, this is so. Moreover, in this case, the metaphysical specification offered in 2 is also plausibly consistent with the standing conception of causation. Without the alleged inconsistencies, the reply fails. The alternative specifications of the Chitra case and of the subset account show that there are no such inconsistencies. Thus, the reply fails.

The second problem with the pre-philosophical conception reply is that there is a tension between its two steps. The first step offers a concrete case, and it succeeds only if we're inclined to say that C in the case caused E2. The second step models this case, and it succeeds only if it successfully models the case by positing downward causation—that is, by positing that C causes E1 and E1 non-causally determines E2. The tension is between the success conditions for the two steps. The first step needs us to accept that C causes E2; the second step asks us to hold onto this claim while accepting a model according to which E2 results from a *non-causal relation*. But accepting that E2 is caused motivates denying that it is brought about

non-causally; and, accepting that E2 is brought about non-causally motivates denying that it was caused. And, importantly, these motivations aren't grounded in the philosophical conception of causation. They're rooted in the characterization of the Downward Causation View itself. No matter how one conceives of causation, it's counterintuitive to say that E2 is both causally and non-causally determined thanks to E1. The reply at hand needs us to overcome such intuitions and accept that the only acceptable model for C's causing E2 is one in which E2 is non-causally determined by one of C's effects. This might not be so bad if the Downward Causation View were the only model available. But it's not. We can also model the case as we did above. That is, we can say that C causes both E1 and E2 immediately. If one is inclined to hold onto the intuition that C causes E2, it's much easier on a model like this than on a model that requires thinking both that (i) E1 non-causally determines E2 and (ii) E1 is the last link in the causal determination of E2.

4. Conclusion

With regard to mental causation, the appearance is that some thoughts cause others. A philosophical account of mental causation ought to explain how this occurs or else 'explain away' the appearance. An account of mental causation that many philosophers seem to think is promising, the Downward Causation View, does neither. Indeed, its commitments regarding the familiar model of non-reductive mental causation seem to be in tension. Drawing on the M1-P2-M2 'diagram' of mental causation, it tells us that (i) P2 is the physical ground, supervenience base, etc. on which M2 depends for its existence; (ii) $P2 \neq M2$; (iii) P2 non-causally determines M2; and, (iv) M1 causes M2 by causing P2. We've asked what justifies claim iv, given that M1 isn't an immediate cause of M2. One might say that iv is true because there is a chain of causal relations from M1 to M2; but this is inconsistent with iii. Alternatively, one might argue that iv is true because P2 and M2 are one and the same effect of M1; but this is inconsistent with ii. As such, an advocate of the Downward Causation View must convince us to reject the common supposition that some C causes E only if C causes E immediately or there is a causal chain from C to E. We considered the possibility that accepting downward causation is more intuitive than rejecting it, but it's implausible. While

examples like Chitra's painting both red and scarlet are intuitive, there's little reason to think that the Downward Causation View is the most intuitive metaphysical specification of them.

Without compelling argument to the contrary, then, mental phenomena seem to be uncaused causes on the Downward Causation View. Advocates of this view seem to be wholly unaware of this. They make no effort to support the replies considered above, nor do they try to mitigate the strangeness of positing so many uncaused causes.¹²

This is a shortcoming of the Downward Causation View; and, it is a shortcoming that some of its competitors do not share. There are reductive alternatives, of course, but there are non-reductive alternatives as well. *Prima facie*, one may avoid the difficulties raised here simply by accepting that mental-to-mental causal relations are immediate. As noted above, various non-reductive views show how this may be done without rejecting physicalism. One may appeal to difference-making or to causal proportionality, for instance. Since these views do not suffer from the shortcomings noted here, there is reason to prefer their non-reductive accounts of mental causation to the Downward Causation View.

For many philosophers, the attraction of downward causation is that it is sensitive to mental-physical dependence. The commitment to physicalism seems to suggest that if there are going to be any mental changes in the world, they must result from physical changes. In an analogy used by Jaegwon Kim (1998, 43), mental phenomena depend on physical phenomena in much the way that aesthetic properties of a painting depend on its physical constitution: if you're to change the painting's aesthetic properties, you have to act on its physical constitution. You can't affect aesthetic properties immediately. If the analogy is apt, then it would seem that you can't affect mental phenomena immediately either. If this is so, then we're saddled with the Downward Causation View. I've argued that this view is inconsistent with our conception of causation; if so, then perhaps we should accept that there simply is no mental-mental causation. Alternatively, we may reject the standing conception of causation and causal chains, or we may reconsider whether physicalism is inconsistent with immediate mental-mental causation. I've claimed that the Downward Causation View gives us little reason to revise our views on causation. I recommend that advocates of downward causation either devise an alternative conception or reconsider their views on what physicalism entails. More generally, we should all be concerned to decide whether our

¹² See, for instance, Bennett (2003), MacDonald and MacDonald (2006), Wilson (2009).

conception of causation and causal chains is inconsistent with physicalism and mental-mental causation.

NOTES

Bennett, K. (2003) Why the exclusion problems seems intractable, and how, just maybe, to tract it. *Nous*, 37(3), pp. 471-497.

– (2007) Mental causation. *Philosophy Compass*, 2(2), pp. 316-337.

– (2008) Exclusion again. In J. Hohwy and J. Kallestrup, eds., *Being Reduced: New Essays on Reduction, Explanation, and Causation*. Oxford: Oxford University Press.

Block, N. (2003) Do causal powers drain away? *Philosophy and Phenomenological Research*, 67(1), pp. 133-150.

Hall, N. (2000) Causation and the price of transitivity. *Journal of Philosophy*, 97, pages 198-222.

Hitchcock, C. (2001) The intransitivity of causation revealed in equations and graphs. *Journal of Philosophy*, 98, pages 273-299.

Kim, J. (1998). *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*. MIT Press.

– (2007) Causation and mental causation. In B.P. McLaughlin and J. Cohen eds. *Contemporary Debates in Philosophy of Mind*. Oxford: Blackwell Publishing.

Lewis, D. (1973) Causation. *Journal of Philosophy*, 70(17), pages 556-567.

Lowe, E.J. (2008) *Personal Agency*. Oxford: Oxford University Press.

MacDonald, C. and MacDonald, G. (2006) The metaphysics of mental causation. *The Journal of Philosophy*, 103(11), pp. 539-576.

Mackie, J.L. (1974) *The Cement of the Universe*. Oxford: Oxford University Press.

Marras, A. (2007) Kim's supervenience argument and nonreductive physicalism. *Erkenntnis*, 66(3), pp. 305-327.

Melnyk, A. (2008) Can physicalism be non-reductive?, *Philosophy Compass* 3(6), pp. 1281-1296.

Menzies, P. (2003) The causal efficacy of mental states. In Walter, S. and Heckmann, H. eds., *Physicalism and Mental Causation*, pages 195-223. Imprint Academic.

Menzies, P. and List, C. (2010). The causal autonomy of the special sciences. In MacDonald, C., editor, *Emergence in Mind*, pages 108–128. Oxford University Press, Oxford.

Ney, A. (2008) Defining physicalism *Philosophy Compass*, 3(5), pp. 1033-1048.

- (2010) Convergence on the problem of mental causation: Shoemaker’s strategy for (nonreductive?) physicalists, *Philosophical Issues*, 20(1), pp. 438-445.
- O’Connor, T. and Churchill, J.R. (2010). Is nonreductive physicalism viable within a causal powers metaphysic? In MacDonald, C. and MacDonald, G. eds., *Emergence in Mind*, pages 43–60. Oxford University Press, Oxford.
- Pereboom, D. (2002). Robust nonreductive materialism. *Journal of Philosophy*, XCIX: 499 – 531.
- Robb, D. and Heil, J. (2013) Mental causation, *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2013/entries/mental-causation/>>.
- Searle, J. (1992) *The Rediscovery of Mind*. Cambridge: MIT Press.
- Shoemaker, S. (2001) Realization and mental causation. In *Identity, Cause, and Mind: Expanded Edition*. Clarendon Press.
- (2013) Physical realization without preemption. In *Mental Causation and Ontology*, S.C. Gibb, E.J. Lowe, and R. D. Ingthorsson (eds.) Oxford: Oxford University Press.
- Sider, T. (2003). What’s so bad about overdetermination? *Philosophy and Phenomenological Research*, 67:719–726.
- Stoljar, D. (2009) Physicalism, *The Stanford Encyclopedia of Philosophy* (Fall 2009 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2009/entries/physicalism/>>.
- Walter, S. (2008) The supervenience argument, overdetermination, and causal drainage: assessing Kim’s master argument. *Philosophical Psychology*, 21(5), pp. 673-696.
- Wilson, J. (1999) ‘How superduper does a physicalist supervenience need to be?’ *The Philosophical Quarterly*, 50: 33–52.
- (2009) Determination, realization, and mental causation. *Philosophical Studies*, 145, pp. 149-169.
- (2011). Non-reductive realization and the powers-based subset strategy. *The Monist* 94: 121-154.
- Witmer, D. G. (2003). Functionalism and causal exclusion. *Pacific Philosophical Quarterly*, 84(2):198–215.
- Woodward, J. (2011), Interventionism and causal exclusion, <http://philsci-archive.pitt.edu/8651>
- Yablo, S. (1992). Mental causation. *The Philosophical Review*, 101(2):245–280.
- Zhong, L. (2011) Can counterfactuals solve the exclusion problem? *Philosophy and Phenomenological Research* 83(1): 129-147.