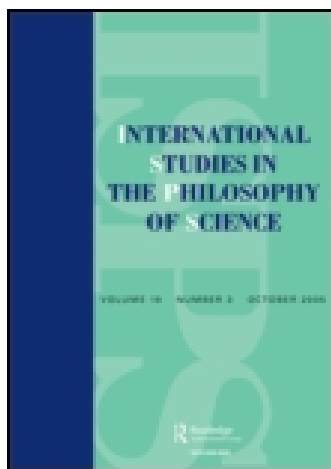


This article was downloaded by: [KU Leuven University Library]

On: 18 September 2014, At: 05:13

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## International Studies in the Philosophy of Science

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/cisp20>

### Interventionism and Supervenience: A New Problem and Provisional Solution

Markus I. Eronen & Daniel S. Brooks

Published online: 12 Sep 2014.

To cite this article: Markus I. Eronen & Daniel S. Brooks (2014) Interventionism and Supervenience: A New Problem and Provisional Solution, *International Studies in the Philosophy of Science*, 28:2, 185-202, DOI: [10.1080/02698595.2014.932529](https://doi.org/10.1080/02698595.2014.932529)

To link to this article: <http://dx.doi.org/10.1080/02698595.2014.932529>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# Interventionism and Supervenience: A New Problem and Provisional Solution

Markus I. Eronen and Daniel S. Brooks

*The causal exclusion argument suggests that mental causes are excluded in favour of the underlying physical causes that do all the causal work. Recently, a debate has emerged concerning the possibility of avoiding this conclusion by adopting Woodward's interventionist theory of causation. Both proponents and opponents of the interventionist solution crucially rely on the notion of supervenience when formulating their positions. In this article, we consider the relation between interventionism and supervenience in detail and argue that importing supervenience relations into the interventionist framework is deeply problematic. However, rather than reject interventionist solutions to exclusion wholesale, we wish to propose that the problem lies with the concept of supervenience. This would open the door for a moderate defence of the interventionist solution to the exclusion argument.*

## 1. Introduction

The causal exclusion argument supposedly shows that mental causes are excluded in favour of the underlying physical causes that do all the causal work (e.g. Kim 1998, 2005). Recently, several philosophers have proposed that if we adopt the interventionist account of causation, the exclusion argument no longer works (Shapiro and Sober 2007; Menzies 2008; Woodward 2008a, 2014; Menzies and List 2010; Raatikainen 2010, 2013; Shapiro 2010, 2012). However, Michael Baumgartner (2009, 2010) has suggested that interventionism gives rise to another exclusion problem that very much resembles

---

Markus I. Eronen is at the Centre for Logic and Analytical Philosophy, KU Leuven. Correspondence to: Centrum voor Logica en Analytische Wijsbegeerte, Hoger Instituut voor Wijsbegeerte, KU Leuven, Kardinaal Mercierplein 2—bus 3200, B-3000 Leuven, Belgium. E-mail: [markus.eronen@hiw.kuleuven.be](mailto:markus.eronen@hiw.kuleuven.be). Daniel S. Brooks is at the Department of Philosophy, Bielefeld University. Correspondence to: Abteilung Philosophie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany. E-mail: [daniel.brooks@uni-bielefeld.de](mailto:daniel.brooks@uni-bielefeld.de)

the original problem formulated by Jaegwon Kim. Interventionism requires that when we intervene on variable  $X$  with respect to  $Y$  we do not change any other variables that are not on the causal path from  $X$  to  $Y$  but are causes of  $Y$  (Woodward 2003). The problem is that it seems to be impossible to intervene on a mental property without also intervening on its supervenience base, so this requirement is violated, and mental properties are threatened to be excluded as causes of physical effects.

Several authors, including James Woodward himself, have responded to Baumgartner's argument (Eronen 2012; Woodward 2014; Weslake forthcoming). The common thrust of these replies is that the way in which Baumgartner represents the exclusion problem violates implicit or explicit constraints for causal modelling, and for this reason his argument does not carry through. In turn, Baumgartner (2013) has replied to this and pointed out that the various ways of avoiding the interventionist exclusion problem all lead to fundamental difficulties.

At the core of this new debate on the exclusion problem is a significant tension that has not been adequately acknowledged by its participants. This tension centres on the prospects of engaging a primarily *metaphysical* problem (exclusion of mental or higher level causes) with a primarily *epistemological* framework (interventionist theory of causation). This is an important underlying premise that motivates the recent so-called 'evidence-based' approaches to solving the causal exclusion problem, as advocated by the philosophers named above. We identify the locus of this tension in the supervenience relation, which is central to the dilemma created by the causal exclusion problem.<sup>1</sup>

In this article, we will analyse the relationship of supervenience and interventionism and demonstrate in detail why the two are incompatible. Following this, we will also offer some insight into why this incompatibility should not lead to a complete rejection of interventionist-inspired solutions or treatments to causal exclusion. Though considerations of space prohibit a full defence of an interventionist solution to causal exclusion, we will argue that opening the door to such a solution will begin by acknowledging long overdue critical stance on the role of supervenience in formulating the basic structure of the problem of exclusion. We will argue that philosophers need to orient their attention to the role of the supervenience relation in both (a) articulating the issue at stake here, namely causal exclusion, and (b) implicitly, and possibly unduly, constraining the space of legitimate solutions to dealing with this issue. Rather than rejecting out of hand evidence-based approaches to engaging causal exclusion, we should question the implicit validity ascribed to the supervenience relation in expressing the issues arising from the problem at hand.

## 2. Interventionism and the Causal Exclusion Problem

The core idea of interventionism is that causes make a difference for their effects: variable  $X$  is causally relevant to a variable  $Y$  if and only if it is possible to carry out an intervention on  $X$  which changes the value or the probability distribution of  $Y$  (Woodward 2003). More precisely:

(M) A necessary and sufficient condition for  $X$  to be a (type-level) direct cause of  $Y$  with respect to a variable set  $\mathbf{V}$  is that there be a possible intervention on  $X$  that will change  $Y$  or the probability distribution of  $Y$  when one holds fixed at some value all other variables  $Z_i$  in  $\mathbf{V}$ . A necessary and sufficient condition for  $X$  to be a (type-level) contributing cause of  $Y$  with respect to variable set  $\mathbf{V}$  is that (i) there be a directed path from  $X$  to  $Y$  such that each link in this path is a direct causal relationship . . . and that (ii) there be some intervention on  $X$  that will change  $Y$  when all other variables in  $\mathbf{V}$  that are not on this path are fixed at some value. (Woodward 2003, 59)

$I$  is an intervention variable for  $X$  with respect to  $Y$  if and only if  $I$  meets the following conditions:

(IV)

11.  $I$  causes  $X$ .

12.  $I$  acts as a switch for all other variables that cause  $X$ . That is, certain values of  $I$  are such that when  $I$  attains those values,  $X$  ceases to depend on the values of other variables that cause  $X$  and instead depends only on the value taken by  $I$ .

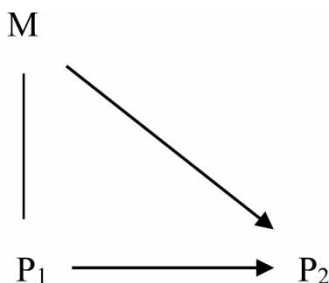
13. Any directed path from  $I$  to  $Y$  goes through  $X$ . That is,  $I$  does not directly cause  $Y$  and is not a cause of any causes of  $Y$  that are distinct from  $X$  except, of course, for those causes of  $Y$ , if any, that are built into the  $I$ – $X$ – $Y$  connection itself; that is, except for (a) any causes of  $Y$  that are effects of  $X$  (i.e., variables that are causally between  $X$  and  $Y$ ) and (b) any causes of  $Y$  that are between  $I$  and  $X$  and have no effect on  $Y$  independently of  $X$ .

14.  $I$  is (statistically) independent of any variable  $Z$  that causes  $Y$  and that is on a directed path that does not go through  $X$ . (Woodward 2003, 98)

It is important for the discussion that follows to note that (M) is defined relative to a variable set  $\mathbf{V}$ , while (IV) is not. If (IV) were also defined relative to a variable set, this would result in a too relativistic and weak account of causation—for example, it would be possible to construct a limited variable set where not all common cause structures are included, and infer entirely spurious causal relations from this set (see Woodward 2008b and Baumgartner 2013 for more). It should be noted, however, that even (M) is not representation relative in any strong sense. As Woodward (2008b) points out, the definition of an intervention guarantees that if  $X$  is a cause of  $Y$  in variable set  $\mathbf{V}$ ,  $X$  will be a (contributing) cause of  $Y$  in any other variable set including  $X$  and  $Y$ . Thus, we can say that if there is a variable set where  $X$  is a cause of  $Y$ , then  $X$  is a cause of  $Y$  *simpliciter*.

The interventionist framework requires that the relata of causation are variables, but states or properties can easily be represented as binary variables (though they do not need to be binary), such that, e.g. value 1 marks the presence of the property and value 0 the absence of the property.

There are several ways of formulating the interventionist exclusion problem; here we adopt a very straightforward rendering (see Baumgartner 2009, 2010, 2013 for details). The problem arises from the following plausible assumptions: (1) interventionism; (2) mental properties supervene non-reductively on physical properties; and (3) mental



**Figure 1** The traditional representation of the causal exclusion problem. Arrows denote causal relations, and the non-directed edge denotes a supervenience relation.

properties are sometimes causes for the physical effects of their supervenience base. This problem is traditionally represented using the schema depicted in [Figure 1](#).

If we then consider (M) and (IV), it follows that in order for mental variable  $M$  to be a cause of physical effect  $P_2$ , there has to be an intervention on  $M$  with regard to  $P_2$  that satisfies the conditions in (IV). However, such an intervention is not possible. Whenever we change  $M$ , supervenience guarantees that there will be a change in  $P_1$ , the supervenience base of  $M$ .  $P_1$  is a variable that is a cause of  $P_2$  but not on the causal path from  $M$  to  $P_2$ . Thus, condition I4 (and possibly also I3) of (IV) is violated. Consequently, it is not possible to intervene on  $M$  with regard to  $P_2$ , and  $M$  cannot be a cause of  $P_2$ . In contrast, due to the asymmetry of supervenience, it is possible to intervene on  $P_1$  with respect to  $P_2$ . Thus, it seems that only  $P_1$  can be a cause of  $P_2$ . This generalizes to all mental variables, and hence, non-reductive mental-to-physical causation is incompatible with interventionism.

Note that this problem arises simply from assumptions (1) to (3)—no further metaphysical principles, such as non-overdetermination, physical causal closure, or exclusion, are needed. It seems that the only way to avoid the problem is to give up non-reductive physicalism, or give up (or revise) interventionism.

### 3. Causal Graphs, Sufficiency, and the Causal Markov Condition

Several authors have recently responded to this argument (Eronen 2012; Woodward 2014; Weslake forthcoming). The common idea behind these responses can be summarized as follows: the representation of mental causation that Baumgartner adopts from Kim is not the kind of causal representation to which interventionism applies, and therefore cannot be used to support an exclusionist conclusion. All three authors appeal to a condition that causal representations are implicitly or explicitly assumed to satisfy: independent fixability (Woodward), independent manipulability (Weslake), or the causal Markov condition (Eronen). Here, we will further explore the approach based on the causal Markov condition, since it reveals interesting relationships between supervenience, interventionism, and latent common causes, but the main conclusions of this article can also be reached by appealing to conditions such as independent fixability or independent manipulability.

As Eronen (2012) argues, the representation of causal exclusion that Baumgartner adopts from Kim violates the causal Markov condition, and is therefore unsuitable for causal modelling. The causal Markov condition (hereafter CMC) is typically stated as follows: conditional on its direct causes, every variable in  $\mathbf{V}$  is independent of every other variable, except its effects (see, e.g. Hausman and Woodward 2004). CMC guarantees that all the (probabilistic) dependencies in the model are due to the causal relationships between the variables in the model. If this condition is violated, making the sorts of interventions required by (IV) becomes impossible, since we cannot hold fixed all the off-path variables. In this sense, CMC is integral to causal modelling, and an implicit assumption underlying interventionism (Hausman and Woodward 2004).

Kim-style representations of mental causation such as Figure 1 clearly violate CMC. Because of supervenience,  $M$  is non-causally correlated with  $P$ : whenever there is a change in  $M$ , there is a change in  $P$ , and when the value of  $P$  is fixed,  $M$  is fixed as well. Thus, conditional on its direct causes,  $M$  is *not* independent of every other variable except its effects.

In response to this, Baumgartner (2013) has argued that CMC can be required only if we *first* assume that the variables of the representation in question are causally sufficient. Causal sufficiency is usually defined as follows: A variable set  $\mathbf{V}$  is causally sufficient if and only if any common cause  $C$  of two variables  $X$  and  $Y$  in  $\mathbf{V}$  belongs to  $\mathbf{V}$ .<sup>2</sup> In the case of the representation in Figure 1, Baumgartner claims, it is far from obvious whether causal sufficiency is satisfied: there may be a common cause of  $M$  and  $P$  that is not included in the variable set.

To properly respond to this point, we need to consider the role of causal sufficiency and CMC in more detail. Causal sufficiency and CMC are conditions that a given set of variables, and their interactions, first need to fulfil before it makes sense to analyse the system in question. The reason that conditions like these are required is that they constitute tools with which to construct representations of systems about which we wish to make causal judgments. In analysing interventionist claims about causal relations, one must pay attention to the causal modelling tradition on which the account is constructed—scientists (and philosophers) cannot expect causal information about a system to simply ‘reveal itself’ to the researcher, but instead a great deal of interpretation is required to reconstruct the workings of a system in a reliable and accurate way. This is a vital component of interventionist theory that cannot be ignored if we wish to evaluate the impact of interventionism on philosophical questions concerning causality. Indeed, it is one of the strengths of interventionism that it is embedded in an account of how we infer causal relations from a set of data, and is what makes the interventionist account ‘closer to scientific practice’ (cf. Shapiro and Sober 2007; Waters 2007, esp. 555; Woodward 2014). What this means for the current discussion is that if the reasoning process behind a particular representation cannot be reconstructed, then this is the reason to doubt the validity of that representation as a suitable case for analysis.

When confronted with data about a system about which we wish to make causal judgements, active interpretation by researchers is necessary. It is not enough to

identify and characterize variables of interest and their (causal) relations to one another, since an indefinite number of representations can be consistent with the data. More importantly, when a representation has been constructed, there are a number of issues that may arise when interventions are introduced to test the causal relations that the representation postulates. For instance, interventions may fail to demonstrate the causal relevance of a particular variable when two variables covary in their changes following an intervention. This may be, for example, due to definitional or some other non-causal dependency between the variables. This is not cause for despair, however, but rather a reason to re-evaluate the description of the system that is supposed to be captured in the representation in question. A case in point is the discussion by Peter Spirtes and Richard Scheines of the causal relation between total cholesterol (TC) and heart disease (HD) (Spirtes and Scheines 2004). Due to a misdescription in the experimental variable (TC), researchers were unable to infer any stable causal relation between TC and HD by intervening on TC. The reason for this is that the variable representing TC was not capable of distinguishing between two particular kinds of cholesterol, LDL (low-density lipids) and HDL (high-density lipids), which have, respectively, provocative and preventive causal effects on the presence of HD. Consequently, any intervention that was made on the TC variable was completely ambiguous for testing the causal relevance of TC for HD (Spirtes and Scheines 2004, 843, table 1). The confounding effect was then corrected by reconstructing the causal graph to take into account the effects of the distinct variables that were latently represented by the original TC variable.

The lesson to draw here from Spirtes and Scheines's discussion is that the description of the system under investigation, and more specifically the variables that are characterized within the variable set  $\mathbf{V}$ , is constantly open to refinement as new information becomes available. In the case of the causal relation between TC and HD, ignorance concerning the actual constituency of the variables postulated for the system under investigation led one of the variables (TC) of the system to be supplanted by two distinct variables (LDL and HDL) that more accurately represented the phenomenon being investigated. Such cases abound in scientific research, which only underscores the importance of acknowledging the relevance of factors influencing our causal reasoning when constructing representations of phenomena.

To return to CMC, the idea behind the conditions of causal sufficiency and CMC is that if we discover that there is an apparent dependency (or similarly confounding problem) among the variables of a graph that is not explained by the causal relations represented in that graph, then that graph's structure needs to be modified in some way. For instance, there may be a 'latent' common cause or causes in the graph that are not represented by any of the variables in that graph's variable set in a way that could explain that dependency (Pearl 2000, 62). Alternatively, some of the variables may be related conceptually, mathematically, definitionally or in some other non-causal way (Hausman and Woodward 2004, 846–847). In both sorts of cases, CMC fails.

Baumgartner takes causal sufficiency to be a precondition for CMC or part of the definition of CMC: 'If  $\mathbf{V}$  is causally sufficient, then every variable in  $\mathbf{V}$  is (probabilistically) independent of all its non-effects in  $\mathbf{V}$  conditional on its direct causes in

V' (Baumgartner 2013, 9; Spirtes, Glymour, and Scheines 2000, 29). In contrast, we take causal sufficiency and CMC to be two distinct conditions, following Hausman and Woodward (2004) and Pearl (2000).<sup>3</sup> In Baumgartner's view, CMC is trivially satisfied when there are missing common causes; in our view, missing common causes is one way in which CMC can fail. Of course, we agree that causal sufficiency and CMC are related in the sense that if we assume or establish causal sufficiency, we know that the possible failure of CMC cannot be due to a missing common cause. However, we believe that the failure of CMC carries important information regardless of whether sufficiency has been established: it indicates that the causal representation is in some way incomplete or misconstrued, and should be revised if possible.

In the end, checking conditions such as causal sufficiency and CMC is part of the process of constructing causal models, whereby various graph structures are tested against one another and against the data out of which the graphs were constructed. What results from this testing are better representations of the phenomena we are interested in.

One could argue that CMC is just a convention, and that we for this reason should not draw any interesting conclusions from its failure (cf. the talk of 'innocuous representational conventions' in Baumgartner 2013, 23). We do not deny the status of CMC as a convention of the interventionist framework, but rather embrace it as such. Judea Pearl, whose account of inferring causality provides one foundation for interventionism, admits this point explicitly: the 'Markov assumption is more a convention than an assumption, for it merely defines the granularity of the models we wish to consider as candidates before we begin the search' for the model most consistent with the data (Pearl 2000, 44). What we do deny is that the conventional status of CMC is a reason to reject its relevancy in analysing causal exclusion from an interventionist perspective. Instead, we affirm such conditions as criterial components that need to be taken into account in order to apply the interventionist framework in the first place. Though some may claim that this limits its usefulness, it also emphasizes the strengths of interventionism as a framework with which to engage in causal analysis.<sup>4</sup> Hausman and Woodward summarize this point:

[E]xpectations [following from CMC] may not be logically inviolable but they seem to be highly reliable in actual application, and they seem to be required if one is able to learn about causal relationships from nonexperimental evidence in the absence of detailed background information. (Hausman and Woodward 2004, 856)

In a similar vein, Pearl writes:

By building the Markovian assumption into the definition of complete causal models ... and then relaxing the assumption through latent structures ... we confess our preparedness to miss the discovery of non-Markovian causal models that cannot be described as latent structures. I do not consider this loss to be very serious, because such models—even if any exist in the macroscopic world—would have limited utility as guides to decisions. For example, it is not clear how one would predict the effects of interventions from such a model, save for explicitly listing the effect of every conceivable intervention in advance. (Pearl 2000, 61–62)



Let us then return to the case of the representation in Figure 1. According to our approach, we can first ask whether the representation satisfies CMC, and if it does not, we can then ask whether this is due to a missing common cause or some other reason. Once we discover the reason, we can revise the representation accordingly.

Prima facie, the failure of CMC in Figure 1 cannot be due to latent common causes. Supervenience, at least the mental-to-physical supervenience that is at issue here, is supposed to be a non-causal relation that is due to certain synchronic or constitutive or determinative relations between the properties, and *not* (just) due to common causes. If mental properties supervene on physical properties, their correlation cannot be accounted for by common causes—otherwise mental properties and physical properties would not be *non-causally* correlated. Thus, even if we included all the latent common causes in the representation, there should still be a residual correlation between *M* and *P* that could not be explained by the common causes.

If this is the case, the question whether the representation in Figure 1 is causally sufficient is irrelevant, since supervenience implies that there will *always* remain a non-causal correlation between the properties involved. This would mean that, *pace* Baumgartner, Kim-style representations cannot satisfy CMC and are unsuitable for causal modelling to start with.

#### 4. Interventionism and Supervenience

The above discussion raises the need to consider the relationship between interventionism and supervenience more carefully. Could it be the case that *M* and *P* do have an interventionist common cause, supervenience notwithstanding? Are such common causes compatible with the non-causal correlation that supervenience requires? Is interventionism at all compatible with (representing) supervenience?

As we saw above in connection to Baumgartner's argument, one problem in combining interventionism and supervenience is that it is not possible to hold fixed the supervenience base variables while intervening on the supervenient variables, which creates problems for applying (IV) and (M). However, in a recent article, Woodward (2014) has proposed a plausible way to avoid this problem and to accommodate supervenience relations into interventionist causal models. To account for supervenience relations, he proposes revising the definitions (M) and (IV) in the following way (these formulations are adapted from Baumgartner 2013, 13–14, and Woodward 2014, section 7):

(M\*) *X* is a (type-level) direct cause of *Y* with respect to the variable set *V* iff there possibly exists an (IV\*)-defined intervention on *X* with respect to *Y* such that all other variables in *V* that are not related in terms of supervenience (or definition) to *Y* are held fixed, and the value or the probability distribution of *Y* changes.

*X* is a (type-level) contributing cause of *Y* with respect to the variable set *V* iff (i) there is a directed path from *X* to *Y* such that each link on this path is a direct causal relationship and (ii) there possibly exists an (IV\*)-defined intervention on *X* with respect to *Y* such that all other variables in *V* that are not located on a causal path from *X* to *Y* or on a path from a variable *Z* to *Y*, such that *Z* is related

in terms of supervenience (or definition) to  $X$  or  $Y$ , are held fixed and the value or the probability distribution of  $Y$  changes.

(IV\*)  $I$  is an intervention variable for  $X$  with respect to  $Y$  iff  $I$  satisfies I1, I2, I3\*, and I4\*:

I3\*. Any directed path from  $I$  to  $Y$  goes through  $X$  or through a variable  $Z$  which is related to  $X$  in terms of supervenience (or definition).

I4\*.  $I$  is (statistically) independent of every cause of  $Y$  which is neither located on a path through  $X$  nor on a path through a variable  $Z$  which is related to  $X$  in terms of supervenience (or definition).

The rationale behind these modifications is that the requirement that we need to hold the supervenience base variables fixed when we intervene on the supervenient variables is too strict and unmotivated (Woodward 2014, section 8). Variables that are related to  $X$  or  $Y$  as a matter of definition, supervenience, or in some other non-causal way are not treated as potential confounders in good scientific methodology (Woodward 2014). For example, when we intervene on a psychological state to determine whether it is causally relevant to another psychological state, it is unreasonable and even absurd to require that we need to hold all the underlying brain states fixed.

With these revised definitions, it appears to be possible to include supervenience relations in interventionist causal models. Additionally, Baumgartner's argument can be dealt with without imposing any extra constraints on the variable set (such as independent fixability): When we intervene on  $M$  with respect to  $P_2$ , the fact that the supervenience base  $P_1$  always changes does not violate (IV\*). Therefore, there is an (IV\*) intervention on  $M$  with respect to  $P_2$ , and  $M$  can be an (M\*) cause of  $P_2$ . However, if we consider the situation more carefully, these definitions do not make representing supervenience relations in the interventionist framework unproblematic—quite the contrary.

Informally, the problem is the following. If  $M$  supervenes on  $P$ , any change in  $M$  will result in a change in  $P$ . Thus, any interventionist cause of  $M$  will also be invariably associated with changes in  $P$ : whenever we intervene on the cause  $C$  to change  $M$ , there will be a change in  $P$ . This change cannot come from any variables except  $C$  or its supervenience base, since we hold all those other variables fixed when intervening on  $C$ . Thus,  $C$  is an (M\*) common cause of  $P$  and  $M$  that explains why there is a change in  $P$  whenever there is a change in  $M$ . This means that if we include both the supervenient variable and its supervenience base in the same causal representation, there will always be a common cause that explains why the supervenience base  $P$  changes whenever the supervenient variable  $M$  changes.

Let us then formulate the argument more precisely. Consider a variable set, where we have variables  $M$  and  $P$ , the former representing a property that supervenes on the property represented by the latter. Suppose then that the value of  $M$  changes. In the variable set  $\mathbf{V}$ , there should be some variable—call it  $C$ —that causes this change (if the cause variable is not in the set, we can revise the set to include it—remember that interventionist causation is not representation relative in any strong sense here). Since  $C$  is a cause of  $M$ , there is an (IV\*) intervention on  $C$  with respect to

$M$  when we hold all the other variables (that are not related to  $C$  or  $M$  by supervenience) in the set fixed. However, whenever there is change in  $M$ , there has to be a change in  $P$  as well (due to supervenience). Thus, an ( $IV^*$ ) intervention on  $C$  with respect to  $M$  also results in a change in  $P$ . Under some plausible assumptions,<sup>5</sup> this implies that there is also an ( $IV^*$ ) intervention on  $C$  with respect to  $P$  when all the other variables (that are not related to  $C$  or  $P$  by supervenience) in the variable set are held fixed. Consequently,  $C$  is a cause of  $P$  as well.

This means that whenever  $M$  changes, there is a common cause  $C$  that explains why  $P$  changes as well. Since we can apply the same reasoning to any change in any supervenient variable, this result is entirely general: whenever there is a change in a supervenient variable  $M$ , the corresponding change in the supervenience base variable is fully explained by a common cause. The same reasoning applies even if we allow the supervenience base of  $M$  to consist of several properties represented by distinct variables.

This result is very undesirable. First, it implies that whenever there is a supervenience relation between a property  $X$  and a property  $Y$  and variables representing both  $X$  and  $Y$  are included in the variable set, there is at least one interventionist common cause variable for  $X$  and  $Y$ . Secondly, it implies that the covariation between variables  $X$  and  $Y$  is fully explained by the common cause(s).

If we assume some form of causal realism regarding interventionist causes, as most participants in this debate do, this has some peculiar metaphysical implications. As we briefly mentioned at the end of the previous section, supervenience is generally taken to be a *non-causal* relation of necessitation or determination—if  $X$  supervenes on  $Y$ , then if  $s$  has  $Y$  it is necessary that  $s$  has  $X$ , or  $Y$  determines  $X$ . For example, Kim writes:

As is customary, I take mind–body supervenience to involve the idea of dependence—a sense of dependence that justifies saying that a mental property is instantiated in a given organism at a time because, or in virtue of the fact that, one of its physical ‘base’ properties is instantiated by the organism at that time. [Supervenience], therefore, is not a mere thesis of covariation between mental and physical properties; it includes a claim of existential dependence of the mental on the physical. (Kim 2003, 152)

The relationship between supervenience and common causes has not been extensively discussed in the literature, but it is highly plausible that if the covariation between two properties is fully explained by a common cause, it makes little sense to maintain that the properties are also related by supervenience.<sup>6</sup> Metaphysical considerations aside, if we look at the issue purely from the point of view of causal modelling or causal representation, it is clear that if we have a causal model where the covariation between two variables is fully explained by a common cause, it makes little sense to include a further relation of supervenience between those variables in the model. Such a relation would be entirely superfluous and would play no explanatory role at all. In this sense, interventionism seems to be incompatible with representing supervenience in any substantial way.

A further consequence of this problem is that variables such as  $M$  and  $P$  that represent properties related by supervenience will appear to exhibit an equal explanatory

status, as the causal roles of  $M$  and  $P$  will *appear to be* empirically indistinguishable under certain interventions. This follows directly from allowing supervenience relations to be represented by the interventionist framework. Specifically, recall that for any particular intervention on one of the variables related by supervenience, this will (non-causally) change the other variable as well, such that the values of the two variables will covary perfectly under, and only under, the relevant class of interventions.<sup>7</sup> If, for such a relevant class of interventions, both (supervening and supervenient) variables are completely interchangeable with respect to explaining the (causal) change in the effect variable, then there will be no explanatory difference in citing one variable or the other (see also Pernu 2013).<sup>8</sup> This seems to be an unpalatable option, first, because it implies violating the important premise in the current discussion that the two variables represent truly distinct properties. Secondly, this goes against the expectation that an intervention is characterized to test the causal relevance of one particular variable *and not another*. That is, interventions in an experimental context will be constructed specifically to test only the variable for which it is intended to test: The degree to which this intention does not hold when actually implementing experimental interventions (because there is an unintended or unexpected covariational change in another variable) is actually an impetus for further revisions of our understanding of the system being investigated. This was the case with the TC example discussed above. In any case, it would simply be a highly questionable scientific methodology if the experimental variable of interest could be swapped out for another variable that is completely unrelated to the intervention tailored to elicit a causal effect in a certain way. Clearly, something has gone wrong in the discussion, a topic we will return to below.

The main point that we have argued for in this section is that, from an interventionist perspective, including variables representing properties related by supervenience in the variable set leads to unacceptable outcomes. It implies that the dependency between the variables can be fully explained by common causes and that the variables involved are entirely interchangeable under a certain class of interventions. Interventionism seems to be incompatible with representing supervenience relations, or to formulate the problem the other way around, importing supervenience relations into the interventionist framework seems to be misguided.

## 5. Causal Exclusion Revisited

The discussion in the preceding two sections leads to the following conclusion regarding traditional representations such as those in Figure 1. If the supervenience relation between  $M$  and  $P$  is not fully accounted for by common causes, there is a non-causal correlation between  $M$  and  $P$ , and CMC is violated. If the covariation between  $M$  and  $P$  is fully explained by common causes, there is no non-causal correlation, and consequently no representation of supervenience. As we have argued above, it seems that the latter is the case, but either way, it is not possible to represent supervenience in any coherent or substantial way in (the current form of) interventionism.

This creates a problem for both exclusionists like Baumgartner and the proponents of the interventionist solution to the exclusion problem. The representation of mental causation that is used as a starting point in the debate either violates CMC or does not represent supervenience, and consequently fails to be an accurate or acceptable causal representation of mental-to-physical causation. Furthermore, our conclusion that supervenience can be explained away by means of identifying a common cause entails the frustrating conclusion that we are required to ascribe an equal explanatory status to apparently distinct variables such as  $M$  and  $P$  (regarding a certain class of interventions). Therefore, we cannot conclude anything concerning mental causation based on arguments that take such representations as a starting point. In a way, this relocates the problem: It is not that the combination of interventionism and mental causation is problematic, it is the combination of interventionism and supervenience that is problematic.

We believe that a plausible solution is to restrict the domain of application of interventionism to sets that have no non-causal relationships among the variables (and that consequently satisfy CMC when there are no latent common causes).<sup>9</sup> Indeed, as Woodward (2014) points out, the original account in Woodward (2003) was implicitly intended to apply only to such sets. As we have seen above, trying to include supervenience relations in the variable set leads to problems. Therefore, instead of changing the definition of (M) to (M\*), we should consider (M) to apply only to variable sets where there are no non-causal dependencies, such as supervenience.

However, we still need to adopt (IV\*) instead of (IV). The original definition of an intervention (IV) leads to the exclusion problem pointed out by Baumgartner, even if we restrict  $M$  to strictly causal variable sets. The definition (IV) is not relativized to a variable set, so even if we do not include the supervenience base variables in the variable set, interventions on mental variables with respect to physical effects are not possible. If we adopt (IV\*), the supervenience base variables need not be held fixed when intervening on the supervenient variables, and the problem of exclusion can be avoided.

Thus, with respect to treating mental causation within an interventionist framework, instead of trying to include supervenience relations in causal models, we propose restricting the domain of application of interventionism to sets without non-causal dependencies. With this approach, interventionism and mental causation are at least compatible, since it is clearly possible that there are strictly causal variable sets where mental variables (e.g.  $M$ ) are causes of physical variables (e.g.  $P_2$ ). Since we adopt the (IV\*) definition of an intervention, we need not hold fixed the supervenience base of  $M$  when intervening on  $M$  with respect to  $P_2$ . If there is a variable set where  $M$  is a cause of  $P_2$ , then  $M$  is a cause of  $P_2$  *simpliciter*, as pointed out in section 2. Thus, it is possible that variables representing mental properties are causes for variables representing physical properties.

However, this leads to a possible further problem, pointed out by Baumgartner (2013): it makes downward causation ubiquitous.<sup>10</sup> In other words, it is not possible to have epiphenomenalist causal structures of the following kind:  $M$  supervenes on  $P_1$ ,

$P_1$  is cause of  $P_2$ , and  $M$  is *not* a cause of  $P_2$ . The (IV\*) definition guarantees that there will always be a variable set (e.g. including only  $M$  and  $P_2$ ) where  $M$  is cause of  $P_2$ . In our view, however, this is a feature and not a bug. It is important to remember that in the interventionist framework causation is not considered to be a metaphysical relation of producing or bringing about the effect; causation is a matter of difference-making and potential manipulation and control. We can make a difference on the value of  $P_2$  by manipulating  $M$ , and we can make a difference on the value of  $P_2$  by manipulating  $P_1$ . There is nothing contradictory or inconsistent about this. It should also be noted that epiphenomenalist causal structures where, for example,  $M$  supervenes on  $P_1$  but is not a cause for anything are very strange and have been traditionally considered as something to avoid. Thus, it is a desirable consequence that interventionism rules out such structures.

We find this approach to mental causation coherent and scientifically plausible. It shows that (*contra* Baumgartner) interventionism is compatible with mental-to-physical or downward causation. This solution may be too weak to satisfy more metaphysically oriented philosophers of mind, and in any case, it does not constitute a ‘silver bullet’ solution to the problem of causal exclusion—the best we can conclude at this stage is that interventionism and mental causation are compatible. However, for all non-metaphysical purposes this is entirely sufficient.

All in all, there remains much to be said about the distribution of philosophical labour concerning how to understand, and analyse, the problem of causal exclusion. Particularly, it is unclear what an evidence-based approach to causality, and specifically the literature on causal reasoning that informs this approach, can contribute to a problem that has hitherto been articulated and subsequently treated, almost exclusively, by traditional metaphysics. Interventionism has been advertised as a theory of causality that is more nuanced than competing philosophical approaches to causality, in large part because of its purported proximity to actual scientific reasoning about causal relations in phenomena that are by all accounts *complex*. Hence, our conclusion that interventionism and mental causation are *at least compatible* should not be dismissed as simply too weak. Rather, it opens the floor for a more nuanced approach to a long-argued issue that preoccupies researchers far outside of metaphysics. This necessitates at least acknowledging the adage often seen in the philosophy of science: *it's more complicated than that!*

## 6. Concluding Remarks

To conclude, we address a possible objection to our general approach, concerning the coherence of analysing supervenience within an evidence-based approach. Being a metaphysical notion, there may be no way of detecting supervenience between two variables by empirical means. If this is the case, it could be argued that it is simply *ad hoc* for us to criticize supervenience from an evidence-based perspective, since there can be no kind of empirical evidence for supervenience. Furthermore, if, as argued above, supervenience in the interventionist framework is completely accounted for by a common cause between the  $M$  and  $P$  variables, it appears to be entirely

superfluous for us to refer *additionally* to supervenience, instead of simply accepting that there *is no* supervenience between *M* and *P*.

We have two responses to offer to this criticism. First, it is simply false to claim that by working from an evidence-based perspective we are thereby blocked from referring to non-empirical concepts or ideas. This, indeed, was one of the lessons in our discussion in section 3: In the face of problematic results where, e.g. two distinct variables exhibit covariational changes under otherwise well-defined interventions, this is reason enough to reconsider the structure of the system that is being represented. Modifying or refining our representations of the system of interest in such circumstances will often require one to postulate *further* relations *outside of* the formalistic aspects of interventionism and its embedding causal modelling framework (in addition to more empirical investigation). Even though they are not explicitly modelled, these relations may, for example, constrain the range of models available.<sup>11</sup> If scientists, or for that matter empirically minded philosophers, had recourse *only* to empirical data and methods to interpret their findings or refine their initial or prior descriptions, scientific progress would essentially cease to exist.

Secondly, and more basically, we take it as inherently justified to speak of supervenience here, at least initially, because its coherence is a basic assumption in the recent arguments for an evidence-based solution to the problem of causal exclusion (cf. Shapiro and Sober 2007; Menzies 2008; Woodward 2008b, 2014; Raatikainen 2010; Shapiro 2010, 2012). Dealing with the supervenience between mental and physical properties is guaranteed in such a discussion because it constitutes a central premise of the causal exclusion problem itself. Assuming that properties *M* and *P* are not identical, and given the widely acknowledged lack of consensus among philosophers regarding how to articulate the exact relation between *M* and *P*, supervenience has become the primary concept for expressing the central issue at hand in the current debate, namely the curious relation between the mental and the physical. The purpose of our analysis in this article, in fact, has been to show that there are grievous problems with trying to import the notion of supervenience wholesale into the interventionist framework.

We thus recognize the pertinence of this objection, particularly since this debate concerns a basic tension between metaphysical and evidence-based perspectives on causal exclusion. However, we believe it is far too early in the game to draw far-reaching consequences about who is allowed to refer to supervenience and in what way. One purposes of this article has been to *initiate* an investigation concerning what prospects there are *at all* for dealing with a metaphysical problem (including its correspondingly metaphysical terminology and concepts) within an evidence-based epistemological (though not metaphysically neutral) framework. Such an investigation has, until now, been lacking on both sides of the debate. The initial answer to this question, perhaps unsurprisingly, is that this will not proceed in a straightforward way. The more substantive answer, as we hope this article to have shown, is that the problem lies with the concept of supervenience itself.

The results of this article force participants on both sides of the debate to reconsider their positions. In particular, they force us to reconsider the way that philosophers are

expected to articulate, and evaluate, the causal exclusion problem. The placative manner in which supervenience is expected to hold, coupled with the way that it figures into the structure of the causal exclusion problem, compels an unwarranted critical evaluation of any interventionist-inspired treatment or solution to exclusionist worries. As we have argued, acknowledging this could open the door for evidence-based solutions to the causal exclusion problem, including interventionist-inspired solutions. Further elaborating such a defence would require exploring in more detail the consequences of the acknowledged incompatibility of interventionism and supervenience. We believe that these consequences will be a source of innovative insight by allowing other concurrent considerations to contribute towards formulating the debate. One such consideration includes what exactly we want from a good theory of causation.

### Acknowledgements

Many thanks to Michael Baumgartner, Laura Bringmann, Lena Kästner, Marie Kaiser, Beate Krickel, and two anonymous referees of this journal for useful and constructive feedback. Earlier versions of this paper were presented at GWP2013 (Hanover, Germany, March 2013), CLPS13 (Gent, Belgium, September 2013), and the Biology discussion group at Bielefeld University; we thank the audiences of these events for their helpful comments. This study has been financially supported by The Research Foundation–Flanders (FWO) (Postdoctoral fellowship, Markus Eronen).

### Notes

- [1] Of course, it is possible to formulate causal exclusion problems without referring to supervenience. Nonetheless, supervenience is, as a matter of fact, a central premise used in mainstream formulations of causal exclusion (see, e.g. Kim 1998, 30). The appeal of supervenience in this regard is easy to understand, because it offers a means of postulating a relatively neutral (ontologically speaking), non-identity relation between different properties whose causal efficacy seem to be in tension with one another. Even when supervenience is not explicitly included in the arguments, it is a background assumption, which is enough to result in the problems we discuss in this paper. Thus, our main theses are also relevant for versions of the causal exclusion argument formulated without supervenience.
- [2] This is the standard definition, but to be exact, sufficiency should be defined as follows: a set  $V$  is causally sufficient iff any common cause of two variables  $X$  and  $Y$  in  $V$  either belongs to  $V$  or has a cause or an effect that is also a common cause of  $X$  and  $Y$  that belongs to  $V$  (Baumgartner 2013, 9). This difference has no implications for our arguments.
- [3] However, our points regarding the relation between supervenience and interventionism also hold if we adopt Baumgartner's definition of CMC, so nothing crucial turns on this.
- [4] These considerations indicate a possibly deep tension underlying our respective treatments of causal exclusion within the interventionist framework and that of Baumgartner's, and have direct bearing on the points to implicit criteria that structure this discussion. By focusing on the metaphysical aspects of the issue, Baumgartner seems to assume that any commitments to the particular conditions that inform *the means by which we represent or come to know* the relations that we judge to be causally related only distract us from the true enterprise confronting us. We strongly disagree with this, and believe that the attitude underlying



it constitutes an out of hand rejection of any interventionist treatment of causal exclusion. Space prohibits in-depth analysis of this point here, but we will elaborate on it in a follow-up article that is in preparation.

- [5] If  $I$  is an intervention variable on  $C$  with respect to  $M$ , it follows from (IV\*) that  $I$  can still fail to be an intervention variable on  $C$  with respect to  $P$  if (1) it is a direct cause of  $P$ , (2) it is a cause of some other variable  $Z$  that is distinct from  $C$  and not on the causal path from  $C$  to  $P$  and is not related by supervenience to  $C$ , or (3) it is not statistically independent of a variable  $Z$  that causes  $P$  and that is on a causal path that does not go through  $C$  or any variable that is related to  $C$  by supervenience. Options (2) and (3) are possible only if the confounding variables are not included in the variable set  $\mathbf{V}$ , since it is assumed that all variables in  $\mathbf{V}$  that are not  $C$  or  $M$  or related to  $C$  or  $M$  by supervenience are held fixed. However, the supervenience relationship between  $M$  and  $P$  guarantees that even if we included all causes of  $M$  and  $P$  in  $\mathbf{V}$ , and held all of them (except  $C$  and its supervenience base) fixed, there would still be a change in  $P$  whenever we intervene on  $C$  with respect to  $M$ . Option (1) also seems very implausible—it is difficult to see how  $I$  could be a cause of  $C$ , which causes the change in  $M$ , and at the same be a direct cause of the change in  $M$ 's supervenience base. In any case, even if  $I$  fails to be an intervention variable for  $C$  with respect to  $P$  for this reason, this is not a problem for our argument: If  $I$  is direct cause of  $P$  and a contributing (via  $C$ ) cause of  $M$ , the dependency between  $M$  and  $P$  is still explained by a common cause (in this case,  $I$ ).
- [6] This exhibits quite well the central problem we pointed out in the introduction: The tension between the largely metaphysical attitude involved in the traditional debate about causal exclusion, and supervenience more generally, and the largely epistemological attitude involved in much of the literature surrounding interventionism.
- [7] The relevant class of interventions would be those interventions on the causal variable that are capable of making a change in the effect variable per the conditions of (IV\*). This point here is simply to exclude interventions that change the causal variable but are not sufficient to bring about a change in the effect variable, as in the case of changing the position of the light switch in a way that does not turn on the light (Woodward 2003, 66–67). In other words, the relevant class of interventions should designate a contrast class of values that the causal variable can take in order to elicit its relevance to causing a change in the effect variable. This point bears directly to variables related by supervenience. Namely, with non-identical supervenient pairs of properties, there can be changes to the supervenience-base property that will not change the supervenient property. However, the reverse will not hold, i.e. any change to the supervening property will lead to a change in the base property. Hence, it is important to be clear that we are only talking here about those intervention-induced changes that preserves the dependency expressed by supervenience in the sense that a change implemented by an intervention will (a) cause a change in the effect variable and (b) exhibit covariant change in both the variables related by supervenience.
- [8] Background assumptions or commitments may give further information for why one variable may be more desirable than the other for a particular explanation that is offered (a desire for explanatory parsimony, or in defense of a more unified explanation), but this seems to avoid the issue at hand. More importantly, this manoeuvre can go in both directions, as other background assumptions or commitments may equally support the other variable.
- [9] Eronen (2012) has briefly proposed this kind of solution, and Weslake (forthcoming) defends a similar approach in a sophisticated formal framework.
- [10] Baumgartner (2013) also points out another potential problem for revised versions of interventionism: we can always include an intermediate variable  $P'$  between  $M$  and  $P_2$ , and then another intermediate variable  $P''$  between  $M$  and  $P'$ , and so on, so that  $M$  directly causes only the first physical event type outside its own supervenience base. However, this objection

requires giving direct causes (as opposed to contributing causes) a special metaphysical status, which is something Woodward (2008b) explicitly denies. Even if we add intermediate variables,  $M$  remains a contributing cause for  $P_2$ .

- [11] For instance, definitional, constitutional, and compositional relations between variables are all consistent with appearing as common cause structures from within the interventionist framework. For that matter, these relations are also all consistently representable with supervenience. Each of these relations nonetheless designate very distinct ways of relating two variables, and it is hence difficult to imagine that any serious researcher would be satisfied with two covarying variables simply being related by a common cause and moving on. This is an issue we will discuss in more detail in a follow-up article that is in preparation.

## References

- Baumgartner, Michael. 2009. "Interventionist Causal Exclusion and Non-reductive Physicalism." *International Studies in the Philosophy of Science* 23: 161–178.
- Baumgartner, Michael. 2010. "Interventionism and Epiphenomenalism." *Canadian Journal of Philosophy* 40: 359–383.
- Baumgartner, Michael. 2013. "Rendering Interventionism and Non-reductive Physicalism Compatible." *Dialectica* 67: 1–27.
- Eronen, Markus. 2012. "Pluralistic Physicalism and the Causal Exclusion Argument." *European Journal for Philosophy of Science* 2: 219–232.
- Hausman, Daniel, and James Woodward. 2004. "Manipulation and the Causal Markov Condition." *Philosophy of Science* 71: 846–856.
- Kim, Jaegwon. 1998. *Mind in a Physical World*. Cambridge, MA: MIT Press.
- Kim, Jaegwon. 2003. "Blocking Causal Drainage and Other Maintenance Chores with Mental Causation." *Philosophy and Phenomenological Research* 67: 151–176.
- Kim, Jaegwon. 2005. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press.
- Menzies, Peter. 2008. "The Exclusion Problem, the Determination Relation, and Contrastive Causation." In *Being Reduced: New Essays on Reduction, Explanation, and Causation*, edited by Jakob Hohwy and Jesper Kallestrup, 196–217. Oxford: Oxford University Press.
- Menzies, Peter, and Christian List. 2010. "The Causal Autonomy of the Special Sciences." In *Emergence in Mind*, edited by Cynthia Macdonald and Graham Macdonald, 108–128. Oxford: Oxford University Press.
- Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
- Pernu, Tuomas. 2013. "Does the Interventionist Notion of Causation Deliver Us from the Fear of Epiphenomenalism?" *International Studies in the Philosophy of Science* 27: 157–172.
- Raatikainen, Panu. 2010. "Causation, Exclusion, and the Special Sciences." *Erkenntnis* 73: 349–363.
- Raatikainen, Panu. 2013. "Can the Mental Be Causally Efficacious?" In *Regarding the Mind, Naturally: Naturalist Approaches to the Sciences of the Mental*, edited by Konrad Talmont-Kaminski and Marcin Milkowski, 138–166. Newcastle: Cambridge Scholars Publishing.
- Shapiro, Lawrence A. 2010. "Lessons from Causal Exclusion." *Philosophy and Phenomenological Research* 81: 594–604.
- Shapiro, Lawrence A. 2012. "Mental Manipulations and the Problem of Causal Exclusion." *Australasian Journal of Philosophy* 90: 507–524.
- Shapiro, Lawrence A., and Elliott Sober. 2007. "Epiphenomenalism: The Dos and the Don'ts." In *Thinking about Causes: From Greek Philosophy to Modern Physics*, edited by Gereon Wolters and Peter Machamer, 235–264. Pittsburgh, PA: University of Pittsburgh Press.
- Spirtes, Peter, Clark Glymour, and Richard Scheines. 2000. *Causation, Prediction, and Search*. New York: Springer.

- Spirtes, Peter, and Richard Scheines. 2004. "Causal Inference of Ambiguous Manipulations." *Philosophy of Science* 71: 833–845.
- Waters, C. Kenneth. 2007. "Causes that Make a Difference." *Journal of Philosophy* 104: 551–579.
- Weslake, Brad. Forthcoming. "Exclusion Excluded." *International Studies in the Philosophy of Science*.
- Woodward, James. 2003. *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.
- Woodward, James. 2008a. "Mental Causation and Neural Mechanisms." In *Being Reduced: New Essays on Reduction, Explanation, and Causation*, edited by Jakob Hohwy and Jesper Kallestrup, 218–262. Oxford: Oxford University Press.
- Woodward, James. 2008b. "Response to Strevens." *Philosophy and Phenomenological Research* 77: 193–212.
- Woodward, James. 2014. "Interventionism and Causal Exclusion." *Philosophy and Phenomenological Research*. doi:10.1111/phpr.12095