Synthese (2024) 203:206
https://doi.org/10.1007/s11229-024-04608-w

ORIGINAL RESEARCH

# Mental causation, interventionism, and probabilistic supervenience

Alexander Gebharter[1] · Maria Sekatskaya[2]

Received: 21 June 2023 / Accepted: 25 April 2024

**Abstract**
Mental causation is notoriously threatened by the causal exclusion argument. A prominent strategy to save mental causation from causal exclusion consists in subscribing to an interventionist account of causation. This move has, however, recently been challenged by several authors. In this paper, we do two things: We (i) develop what we consider to be the strongest version of the interventionist causal exclusion argument currently on the market and (ii) propose a new way how it can in principle be overcome. In particular, we propose to replace strict supervenience in the assumption that the mental supervenes on the physical by probabilistic supervenience and show how this move has the potential to license the inference to mental causation. Finally, we argue that probabilistic supervenience captures some of the most important intuitions that strict supervenience captures and discuss possible objections to weakening strict supervenience in the way we suggest.

**Keywords** Causal exclusion · Interventionism · Mental causation · Supervenience

## 1 Introduction

Intuitively, it seems plausible that one's mental states are not identical to their physical realizers and have some autonomy *vis-à-vis* the latter. Mental states make a difference for how the future unfolds, which allows us to interact with our environment and gives us to some extent control over our lives. However, mental causation is threatened by

✉ Alexander Gebharter
alexander.gebharter@gmail.com

Maria Sekatskaya
maria.sekatskaya@gmail.com

[1] Center for Philosophy, Science, and Policy (CPSP), Department of Biomedical Sciences and Public Health, Marche Polytechnic University, Via Tronto 10A, 60126 Ancona, Italy

[2] Institute of Philosophy, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

Published online: 11 June 2024

several philosophical objections. One of the most prominent is the causal exclusion argument (Kim, 2005). In a nutshell, the argument says that if the mental is ontologically non-identical to the physical, the mental supervenes on the physical, for every physical state there is a sufficient physical cause, and causal overdetermination does not happen systematically, then the mental is causally inert.

Supporters of mental causation have several options to block[1] the exclusion argument. One can either attack (A) one or several of the argument's premises or (B) the argument's validity. One way to contest the truth of the premises or the argument's validity consists in further specifying the notion of causation used in the argument (cf. Hitchcock, 2012). This paper builds on the debate that resulted from one such attempt, in particular, from the attempt to escape causal exclusion by subscribing to an interventionist account of causation such as Woodward's (2003). Some of the advantages of interventionism, if compared to other theories of causation, is that it is close to scientific practice, comes with a clear semantics for causation as well as a sophisticated methodology for causal inference, and provides a rich and easy to expand basis for exploring a multitude of philosophical issues directly related to causation.[2] We do not go into further details why interventionism is an especially promising candidate for reconstructing the causal exclusion argument but rather content ourselves with pointing the reader to the fact that there exists a rich literature about this specific topic and let this fact speak for itself (see, e.g., Baumgartner, 2010, 2013, 2018; Eronen & Brooks, 2014; Gebharter, 2017a; Hoffmann-Kolss, 2014, 2022; Kinney, 2023; Stern & Eva, 2023; Woodward, 2015).

Because much has been written about causal exclusion and interventionism, we need to narrow down the scope of this paper even further. In particular, we will only consider what in our view is the most threatening objection to interventionist mental causation currently on the market. This threat emerged from a series of objections raised by authors such as Baumgartner (2010, 2013), Baumgartner and Gebharter (2016), Gebharter (2017a, 2017b), and culminated in Baumgartner's (2018) "The inherent empirical underdetermination of mental causation". Our main focus will not lie on any particular argument to be found in these papers, but rather on what we take to be the strongest argument based upon these arguments. In a nutshell, the argument we will develop aims at establishing that a mental state causing some other mental or physical state is excluded by the core definitions underlying interventionism.

Classical causal exclusion arguments (e.g., Kim, 2005) are often used as arguments against non-reductive physicalism, which, as any version of physicalism, crucially relies on the assumption of the causal closure of the physical domain that prominently figures as a key premise in these arguments. In a nutshell, causal closure says that for any physical event there is a sufficient physical cause. The interventionist causal exclusion argument we will discuss is stronger because it concludes the causal inefficacy of the mental even without assuming that the physical domain is causally closed. Thus, it threatens not only non-reductive physicalism but any philosophical

[1] By *blocking* the exclusion argument we mean to render mental causation conceptually compatible with the truth of the other premises used in the argument. Thus, blocking the exclusion argument does not yet guarantee that there actually is mental causation. It rather paves the road for that by establishing the possibility of mental causation given the premises are true. This is what the present paper is about.

[2] See, for example, (de Grefte & Gebharter, 2021; Gebharter et al., 2019).

view that assumes mental causation as well as that the mental is non-identical to the physical and supervenes on the physical, regardless of whether said view is a variant of physicalism or dualism. Therefore, we frame the main question to be pursued in this paper as whether and how mental causation is possible from an interventionist perspective given that the mental non-trivially supervenes on the physical and not to the background of the non-reductive physicalism vs. reductive physicalism debate. We believe that we are in good company in doing so and point the reader to other papers on interventionist causal exclusion which are mainly framed in terms of mental causation and epiphenomenalism rather than non-reductive vs. reductive physicalism (e.g., Eronen & Brooks, 2014; Kistler, 2017).

In this paper, we zoom in on one specific assumption required to get the interventionist exclusion argument going: The mental supervenes on the physical. Supervenience is typically understood as a strict relation saying that there is no difference in higher-level supervenient properties without a difference in lower-level subvenient properties. This relation allows non-reductionists to assert the ontological non-identity of the mental and the physical. We argue that strict supervenience can be replaced by probabilistic supervenience in such a way that the interventionist exclusion argument is blocked.[3] Since probabilistic supervenience cannot be ruled out metaphysically, it provides an in-principle rebuff to the exclusion argument. We argue that weakening supervenience in the way we suggest does no harm because it satisfies the same core intuitions underlying the strict standard version of supervenience to which the modern scientific picture of the world typically adheres. We also discuss further possible objections against weakening supervenience and provide independent motivation for this move. This independent motivation suggests that probabilistic supervenience might not only block the interventionist exclusion argument but also open up a whole new area for supporters of mental causation to explore. For example, it seems to allow for a difference between mental and physical causation and to ground the intuition behind the explanatory gap argument.

The paper is structured as follows. In Sect. 2, we introduce the basics of interventionism required for the rest of this paper. In Sect. 3, we develop what we consider to be the strongest version of the interventionist exclusion argument currently on the market. In Sect. 4, we propose to weaken supervenience and argue that this move allows the interventionist to block causal exclusion. In Sect. 5, we finally discuss possible objections to the proposed weakening of supervenience and provide independent motivation for probabilistic supervenience. We conclude in Sect. 6.

## 2 Interventionism

The basic intuition underlying interventionism is that whenever one variable $C$ is causally relevant for another variable $E$, it is in principle possible to change $E$'s

[3] This is a type (B) strategy. It attacks the validity of the argument. In particular, we show that the argument's original conclusion does no longer follow from the premises given a novel and weaker interpretation of the supervenience assumption.

value by manipulating $C$ based on suitable interventions.[4] Interventions do not need to be realizable by human actions. Thus, interventionism is a non-anthropocentric manipulability theory of causation. For our endeavor, it suffices to introduce only a few basic concepts, the first of them being the notion of direct causation (Woodward, 2003, p. 55):

> **(DC)** C is a direct cause of $E$ w.r.t. a set of variables $\mathbf{V}$ if and only if there is a possible intervention on $C$ w.r.t. $E$ that changes $E$'s value while the values of all other variables in $\mathbf{V}$ are held fixed by additional interventions.

Direct causal relationships are represented by directed edges in a causal graph. $C \rightarrow E$ , for example, stands for $C$ being a direct cause of $E$. Concatenations of such causal relations are called causal paths. A causal path of the form $C \rightarrow \ldots \rightarrow E$ is called a directed causal path from $C$ to $E$. Structures consisting only of variables and causal arrows are called causal graphs. If they do not feature causal loops, they are called directed acyclic graphs or DAGs for short. In this paper, we will only consider causal structures that are DAGs.

Based on the notion of direct causation and the notion of a directed causal path, one can define the following notion of contributing causation (Woodward, 2003, p. 59):

> **(CC)** $C$ is a contributing cause of $E$ w.r.t. a set of variables $\mathbf{V}$ if and only if there is a directed causal path $\pi$ from $C$ to $E$ and a possible intervention on $C$ w.r.t. $E$ that changes $E$'s value while the values of all other variables in $\mathbf{V}$ not lying on $\pi$ are held fixed by additional interventions.

This paper is not the place for a detailed motivation of **(DC)** and **(CC)**. We would like to point interested readers to (Woodward, 2003, Sect. 2.3) fur further details.

Note that **(DC)** and **(CC)** are both relativized to a set of variables $\mathbf{V}$. It can happen that a variable $C$ turns out to be a direct or contributing cause of another variable $E$ w.r.t. one variable set $\mathbf{V}$, but not w.r.t. another variable set $\mathbf{V}'$. This can, for example, happen if there are several directed causal paths from $C$ to $E$ that cancel each other out. If the variables lying on these paths are represented in $\mathbf{V}$, but not in $\mathbf{V}'$, then **(DC)** and **(CC)** will be able to account for $C$'s causal relevance for $E$ w.r.t. $\mathbf{V}$, but not w.r.t. $\mathbf{V}'$. This means that whether a causal relation holds according to **(DC)** and **(CC)** depends on the choice of variables. Since this paper is about metaphysics and not methodology, we also need a third causal notion that is not relativized to a particular set of variables (Woodward, 2008, p. 209):

> **(CS)** $C$ is a cause of $E$ (simpliciter) if and only if there is an admissible variable set $\mathbf{V}$ such that $C$ is a contributing cause of $E$ w.r.t. $\mathbf{V}$.

By an admissible variable set we mean a variable set that satisfies the following condition of independent fixability (Woodward, 2015, p. 316):

> **(IF)** $\mathbf{V}$ satisfies independent fixability if and only if any combination of values of variables in $\mathbf{V}$ can be brought about by possible interventions.

[4] For technical reasons, interventionism first and foremost understands causation as a relation that holds between variables. It can, thus, be considered as an account of type causation. By adding further definitions to the theory's core, one can expand interventionism to also cover token or actual causation (see, e.g., Woodward, 2003).

**(IF)** can be expected to regularly be violated in the presence of non-causal relations,[5] but is assumed to hold for any set containing only variables describing distinct events. More precisely, distinctiveness means that for any two variables in such sets, the instantiations of the event-types represented by the values of these variables do not spatiotemporally overlap (cf. Kim, 1973; Lewis, 1986; Woodward, 2016). In order to guarantee that any causal relation output by interventionism does indeed not conflate causation with some other form of dependence, we endorse this assumption throughout the argument: If $C$ is a cause of $E$ (simpliciter), then there is a set of variables describing distinct events such that $C$ is a contributing cause of $E$ w.r.t. that set of variables. We thus follow Woodward's (2015) suggestion that only sets whose variables describe distinct events and can be controlled independently should be used for evaluating causal claims.

Finally, we need to clarify what it means to intervene on a variable $C$ w.r.t. another variable $E$. Only interventions that satisfy certain conditions will be suitable to account for causal relations. These conditions are summarized in the following definition of an intervention variable (based on Woodward, 2003, p. 98):

(IV) $I_C$ is an intervention variable for $C$ w.r.t. $E$ if and only if
(i) $I_C$ is a cause of $C$, and
(ii) if $I_C$ is a cause of $E$, then only via a path going through $C$, and
(iii) $I_C$ is probabilistically independent of any cause $X$ of $E$ that causes $E$ over a path not going through $C$.

The notion of causation used in **(IV)** should be understood as causation (simpliciter) as defined in **(CS)**. Thus, also **(IV)** is not relativized to a particular set of variables (Woodward, 2008).[6] Also note that in Woodward's (2003) original version there is a fourth condition requiring that the intervention variable, if active, decouples $C$ from all its other causes. Since this condition is not required to get the interventionist machinery working (Baumgartner & Drouet, 2013, pp. 186f), we can ignore it in this paper.[7]

Finally, we can define an intervention on $C$ w.r.t. $E$ as an intervention variable $I_C$ for $C$ w.r.t. $E$ taking one of its values $i_C$ such that setting $I_C$ to this value $i_C$ is associated with a change of $C$'s value.

Before we proceed by presenting the interventionist exclusion argument in Sect. 3, let us introduce a slight modification of **(IV)**. In fact, conditions (ii) and (iii) can be weakened. The original versions are intended to guarantee that any change in $E$ associated with an intervention $I_C = i_C$ indicates a causal influence of $C$ on $E$. This can, however, be guaranteed even if $I_C$ causes $E$ also over a path not going through $C$ (which violates condition (ii)) and if $I_C$ is probabilistically dependent on another cause $X$ of $E$ not lying on a directed path from $I_C$ to $E$ going through $C$ (which violates condition (iii)). This is the case if $I_C$'s taking on its value $i_C$ is not associated with any change of values of any off-path variable $X$ that causes $E$, i.e., any variable $X$ that causes $E$ not lying on a directed path from $I_C$ to $E$ going through $C$. To see the

[5] For several cases where **(IF)** is satisfied even in the presence of non-causally dependent variables see, for example, (Hoffmann-Kolss, 2022).

[6] **(IV)** is not relativized to a particular variable set because otherwise interventionism would lead to obviously false causal claims. For details, see (Baumgartner, 2013).

[7] Our condition (i) corresponds to condition I1, (ii) to I3, and (iii) to I4 in (Woodward, 2003, p. 98).

interventionist causal exclusion argument's full power,[8] we use this weakened version of an intervention variable throughout the paper:

(IV)* $I_C$ is an intervention variable for $C$ w.r.t. $E$ if and only if
(i) $I_C$ is a cause of $C$, and
(ii) if $I_C$ is a cause of [9], $i_C$ , then $I_C = i_C$is not associated with any change of values of any cause $X$ of $E$that causes $E$over a path not going through$C$.

We have now all the interventionist basics we need to understand the interventionist causal exclusion argument to be discussed in this paper. We introduce this argument in the next section.

## 3 An interventionist version of the exclusion argument

In this section we introduce what we consider to be the most threatening objection to interventionist mental causation currently on the market. The interventionist exclusion argument we focus on is based on the arguments to be found in (Baumgartner, 2010, 2013, 2018). We will ignore some details and expand upon others to some extent. In particular, our version of the argument even allows the supporter of mental causation to rely on the condition of independent fixability **(IF)** and to endorse the weaker notion of an intervention variable **(IV)*** instead of the original **(IV)**. An advantage of the interventionist exclusion argument is that it does not require all the premises of the original causal exclusion argument. In particular, neither the causal closure of the physical domain nor the assumption that there is no systematic overdetermination is needed. It requires only that the mental is ontologically non-identical to the physical and that it supervenes on the physical. Thus, it not only threatens non-reductive physicalism, but any position relying on non-trivial supervenience of the mental on the physical that wants to uphold mental causation.

The argument comes in the form of a *reductio ad absurdum*. It runs as follows: Assume that the mental variable $M$ is a cause (simpliciter) of another causal variable $E$. This is the assumption that is challenged by the argument by showing that it leads to inconsistencies. Let us further assume that the mental supervenes on the physical, meaning that there will also be a variable $P$ such that any change in $M$-values is associated with a change in $P$-values. From $M$ being a cause (simpliciter) of $P$ it follows with **(CS)** that there is an admissible set of variables $\mathbf{V}$ containing $M$ and $E$ such that $M$ turns out to be a contributing cause of $E$ w.r.t. $\mathbf{V}$. This set will only contain variables representing distinct event types and, thus, it will also satisfy **(IF)**, meaning that all the variables in $\mathbf{V}$ can be independently fixed by interventions. Note that it follows from this that $\mathbf{V}$ does not contain $M$'s supervenience base $P$ since due to supervenience not all combinations of $M$-values and $P$-values describe distinct events and can be brought about independently by interventions. Now for $M$ to be a

[8] As we will see in the next section, the interventionist causal exclusion argument derives its strength from the result that one cannot consistently assume the existence of an intervention variable for any mental variable. Showing that there is not even a weaker intervention variable in the sense of **(IV)*** thus makes the argument even stronger.

[9] $I_C$'s on-values are those values that fix $C$ to a certain value.

contributing cause of $E$ w.r.t. $\mathbf{V}$ there must, according to **(CC)**, exist an intervention $I_M = i_M$ on $M$ w.r.t. $E$ that is associated with a change in $E$ if the values of all off-path variables are fixed by additional interventions. So far so good.

Let us turn to $M$'s supervenience base $P$ now. Since $\mathbf{V}$ is an admissible variable set, also the set $\mathbf{V}'$ that results from $\mathbf{V}$ by replacing $M$ by $P$ will be admissible. The reason for this is that $\mathbf{V}'$ will not contain any pair of variables whose values represent event-types with spatiotemporally overlapping instantiations because also $\mathbf{V}$ did not contain such a pair of variables. To see why this is the case, let us assume that $\mathbf{V}'$ would indeed contain two variables with values representing spatiotemporally overlapping event-types. These two variables are either both different from $P$ or they are not. If they are both different from $P$, then also $\mathbf{V}$ would contain them, which is excluded by $\mathbf{V}$ being an admissible variable set. Thus, one of these variables needs to be $P$. (Let us label the other one $X$.) If this would be the case, then also $M$ would feature values representing event-types that spatiotemporally overlap with event-types represented by values of $X$. This follows from the values of $P$ being the realizers of the values of $M$ and the fact that realizers cannot spatiotemporally exceed what they realize. But also $M$ featuring values representing event-types that spatiotemporally overlap with event-types represented by values of $X$ is already excluded by $\mathbf{V}$ being admissible.

Since we already established that there exists an intervention $I_M = i_M$ on $M$ w.r.t. $E$ that is associated with a change in $E$ when the values of all off-path variables in $\mathbf{V}$ are fixed by additional interventions and since $M$ supervenes on $P$, we can infer that there also exists an intervention $I_P = i_P$ on $P$ w.r.t. $E$ that is associated with a change in $E$ when the values of all off-path variables in $\mathbf{V}'$ are fixed by additional interventions. The reason for this is how $P$-values are mapped onto $M$-values due to supervenience. All that is needed is to be able to bring about one of the $P$-values $p$ realising one of those $M$-values $m$ for which we already established that $I_M = i_M$ is associated with changes in $E$. Since $\mathbf{V}'$ is an admissible variable set, it follows from **(IF)** that there will be such an intervention $I_P = i_P$ on $P$ w.r.t. $E$. This intervention will, since it sets $M$ to $m$ due to the fact that $m$ is realized by $p$, also be associated with changes in $E$. It then follows from **(CC)** that $P$ is a contributing cause of $E$ w.r.t. $\mathbf{V}'$ and from this with **(CS)** that also $P$ is a cause (simpliciter) of $E$.

We have found so far that both $M$ and $P$ are causes (simpliciter) of $E$. Let us now focus on the intervention variable $I_M$. From **(IV)*** it follows that $I_M$ is a cause (simpliciter) of $M$. Thus, from **(CS)** it follows that there must be an admissible variable set $\mathbf{V}''$ containing $I_M$ and $M$ such that $I_M$ is a contributing cause of $M$ w.r.t. $\mathbf{V}''$. From **(CC)** it then follows that there exists an intervention $I_{I_M} = i_{I_M}$ on $I_M$ w.r.t. $M$ such that $I_{I_M} = i_{I_M}$ is associated with a change in $M$ when the values of all off-path variables in $\mathbf{V}''$ are fixed by additional interventions. Let $\mathbf{V}'''$ be the variable set we get from $\mathbf{V}''$ by replacing $M$ by $P$. Again, since $\mathbf{V}''$ was admissible, also $\mathbf{V}'''$ will be admissible (for the very same reasons for which $\mathbf{V}'$ turned out as admissible if $\mathbf{V}$ is). And since $M$ supervenes on $P$ and $I_{I_M} = i_{I_M}$ is associated with a change in $M$ when the values of all off-path variables in $\mathbf{V}''$ are fixed by additional interventions, it will also be associated with a change in $P$ when the values of all off-path variables in $\mathbf{V}'''$ are fixed by additional interventions. The reason for this is that due to supervenience no two different $M$-values can be realised by one and the same $P$-value. So, for any change in $M$-values due to $I_{I_M} = i_{I_M}$ there must be a corresponding change in $P$-values. It then follows with **(CC)** that $I_M$ is a contributing cause of $P$ w.r.t. $\mathbf{V}'''$ and from the latter with **(CS)** that $I_M$ is a cause (simpliciter) of $P$.

Here comes the problem: Since any directed causal path between $M$ and $P$ is excluded by the typical assumption that supervenience is a non-causal relation,[10] it follows that $I_M$ causes $E$ over two different causal paths, one going through $M$ and one going through $P$ such that any intervention $I_M = i_m$ that is associated with changes in $M$-values is, at the same time, associated with changes in $P$-values. This, however, contradicts **(IV)***(ii) and, thus, the earlier result that $I_M$ is an intervention variable for $M$ w.r.t. $E$. Since nothing hinged on the particular choice of the variables we used, this result generalizes: From the assumption that any mental variable $M$ is a cause (simpliciter) of another variable $E$ it follows that there exists an intervention variable $I_M$ for $M$ w.r.t. $E$ and, at the same time, that no such intervention variable exists. Thus, by *reductio* it follows that no mental variable can be a cause (simpliciter) of any other variable whatsoever.

Let us briefly add two comments about this argument: The first one is that the argument works even under the assumption that all the variable sets used in the argument satisfy independent fixability **(IF)**. This is remarkable because several authors (most prominently Woodward, 2015) have argued that interventionist exclusion arguments can be blocked by committing oneself to only using variable sets satisfying **(IF)** for causal evaluations along the lines of **(DC)** and **(CC)**. This is exactly what we did in our version of the interventionist exclusion argument as presented above. We assumed **(IF)** whenever we applied **(DC)** or **(CC)** and only required to hold fixed variables in those specific sets. But as it turned out, not even restricting oneself to sets satisfying **(IF)** does in the end improve the situation for the supporter of mental causation.

The second comment is that Woodward (2015) suggested that the argument (or a version of it) could be blocked by allowing for interventions that violate condition **(IV)***(ii) in cases where supervenience relationships are involved. For our argument this would mean that the consequence that $I_M$ is an intervention variable that influences $M$ as well as $P$ at the same time but over two different causal paths would not amount to a contradiction anymore. However, Baumgartner (2018) showed that in such cases one cannot decide whether the causal work that brings about changes in $E$ is done by the directed path going through $M$, by the directed path going through $P$, or by both paths together. Thus, $M$'s causal efficacy cannot be established conclusively and the threat of epiphenomenalism would still stand. Since the modification to the strict notion of supervenience used in both arguments that we will propose in the next section already blocks the first argument, there is no need to further modify the notion of an intervention variable **(IV)*** and we will stick with the version of the interventionist exclusion argument developed in some detail above.

[10] For a recent proposal to understand supervenience as a causal relation, see (Leuridan, 2012; Leuridan & Lodewyckx, 2020). We point the interested reader to these papers but do not explore the strategy discussed in them further here.

## 4 Probabilistic supervenience to the rescue

In this section we propose to replace the strict notion of supervenience used in the interventionist exclusion argument introduced in Sect. 3 by a weaker probabilistic version. After that, we show how this probabilistic version of supervenience can be used to block the interventionist exclusion argument.

To make our overall strategy more transparent, it is useful to start with a more precise definition of strict supervenience[11]:

**(SUP)**$M$ supervenes on $P$ if and only if (i) each $M$-value is realized by some $P$-value, and (ii) $M$'s value does not vary anymore once $P$'s value is fixed.

Note that **(SUP)** is stated for variables, which makes it directly applicable to the background of an interventionist understanding of causation. (i) guarantees that no $M$-value can occur without a corresponding $P$-value being instantiated at the same time. (ii) reflects the supervenience assumption: Whenever $M$ changes its value, then also $P$ does, or the other way round: If $P$'s value is fixed, then also $M$'s value is.

What caused the problem for mental causation in Sect. 3 was that any cause of $M$ turned out to be a common cause of $M$ and $P$. This is a direct consequence of **(SUP)**. So, what we need to allow for in order to block the interventionist exclusion argument is a notion of supervenience that avoids this consequence. Before we propose such a notion, it will be useful to state **(SUP)** in probabilistic terms:

> **(SUP)***$M$ supervenes on $P$ if and only if (i) each $M$-value is realized by some $P$-value, and (ii) for all $P$-values $p$ there is an $M$-value $m$ such that $Pr(m|p) = 1$.

Now it becomes easy to see how strict supervenience can be weakened. Instead of requiring that every $P$-value $p$ determines $M$ to take one of its values $m$ with probability 1, we only demand that every change in $M$-values makes a probabilistic difference for at least one of $P$'s values. We thus arrive at the following notion of probabilistic supervenience:

> **(SUP)****$M$ supervenes on $P$ if and only if (i) each $M$-value is realized by some $P$-value, and (ii) for all $M$-values $m$ and $m^{'}$ (with $m \neq m^{'}$) there is a $P$-value $p$ such that $Pr(p|m) \neq Pr(p|m')$.

Note that the move from strict to probabilistic supervenience preserves important intuitions about supervenience: Supervenience should anchor or ground the mental in the ongoings in the physical. In other words, supervenience is assumed to restrict what is possible on higher levels given what is happening on the fundamental physical level. This is what makes the supervenience relation so attractive to non-reductionists about mental causation: It allows for the non-identity of physical and mental properties without claiming that mental properties can exist without some fundamental physical properties underlying them.[12] Like strict supervenience, probabilistic supervenience

[11] Prominent definitions of supervenience to be found in the literature often feature necessity operators. (For an overview see, e.g., McLaughlin & Bennett, 2021). These stronger notions imply **(SUP)**. Running the interventionist causal exclusion argument with the weaker version **(SUP)**, as we did in Sec. 3, thus reveals its full strength.

[12] Craver (2017) discusses the same metaphysical relation under the label "stochastic physical supervenience". He concludes that while such a relation is conceptually possible, it is neither well-motivated nor

can do the same job, depending on the specific probabilistic pattern between $M$-values and $P$-values. Once the physical variable $P$ on which the mental variable $M$ supervenes is fixed to a certain value $p$, some mental states might be compatible with $p$, while others might not be compatible.[13] The difference to strict supervenience is that its probabilistic cousin is less restrictive. Now it can happen that more than one mental state $m$ is compatible with a physical state $p$. How restrictive a particular supervenience relation is is, in the end, an empirical question that cannot be answered a priori. Note, however, that also strict supervenience is not inferred from empirical data showing how mental properties actually correlate with physical properties, but is rather accepted on purely metaphysical grounds. If probabilistic supervenience does the same job without facing some of the problems of strict supervenience, this constitutes a good reason to prefer it. In Sect. 5 we will consider different arguments in favor of probabilistic supervenience, as well as some conceptual consequences of accepting it.

Let us now come back to the question of how **(SUP)**** can help us to overcome the interventionist exclusion argument. To this end, we provide a simple model illustrating how $M$'s value can be changed by an intervention variable $I_M$ for $M$ w.r.t. $E$ without changing $P$'s value at the same time. Assume $M$ has two values $m_1, m_2$ and its supervenience base $P$ has four values $p_1$, $p_2$, $p_3$, $p_4$. Further assume that the following conditional probabilities characterize the (probabilistic) supervenience relationship between $M$ and $P$:

$$Pr(m_1|p_1) = 0.75$$

$$Pr(m_1|p_2) = 0.5$$

$$Pr(m_1|p_3) = 0.25$$

$$Pr(m_1|p_4) = 0$$

According to this characterization, $m_1$ can be realised by $p_1$, $p_2$, and $p_3$, while $m_2$ can be realised by all four $P$-values. In particular, the probabilities that each of the individual $P$-values realizes $m_1$ and $m_2$, were $m_1$ and $m_2$ be instantiated, are as follows[14]:

$$Pr(p_1|m_1) = 0.5 \; Pr(p_1|m_2) = 0.1$$

Footnote 12 continued
helpful, although he sees some potential of this relation in addressing the causal exclusion argument. We argue that one should consider probabilistic supervenience because it not only blocks the interventionist exclusion argument, but also offers other interesting empirical and metaphysical possibilities.

[13] Note that $Pr(p|m) \neq Pr(p|m')$ in **(SUP)**** implies that $P$ probabilistically depends on $M$. Since probabilistic dependence is a symmetric relation, it thus further implies that also $M$ probabilistically depends on $P$. Hence, the supervenience base will always have some probabilistic influence on the supervening variable.

[14] For the sake of simplicity, we assume that $Pr(p_i) = 0.25$ for $i \in \{1, 2, 3, 4\}$. The conditional probabilities $Pr(p_i|m_j)$ with $j \in \{1, 2\}$ can be computed with Bayes' theorem as $\frac{Pr(m_j|p_i)Pr(p_i)}{\sum_{p_i,m_j} Pr(m_j|p_i)Pr(p_i)}$.

$$Pr(p_2|m_1) = 0.3\dot{3}\ Pr(p_2|m_2) = 0.2$$

$$Pr(p_3|m_1) = 0.1\dot{6}\ Pr(p_3|m_2) = 0.3$$

$$Pr(p_4|m_1) = 0\ Pr(p_4|m_2) = 0.4$$

Now the interventionist exclusion argument can be blocked as follows: We proceed exactly as in the argument outlined in Sect. 3 until we have established that both the mental variable $M$ and the physical variable $P$ are causes (simpliciter) of the causal variable $E$. The difference is how we proceed from there. In the original argument, we were able to show that $I_M$ is not only a cause simpliciter of $M$, but also of $P$ and, more importantly, that any change brought about in $M$-values due to $I_M = i_M$ corresponded to a change in $P$-values since no two different $M$-values can be realized by one and the same $P$-value due to strict supervenience. But this is not the case if we replace strict by probabilistic supervenience. If we use an intervention $I_M = i_M$ to change $M$'s value from $m_1$ to $m_2$, for example, then $I_M = i_M$ does not necessarily also lead to a change in $P$-values. The reason for this is that both $m_1$ and $m_2$ are compatible with the physical realisers $p_1$, $p_2$, and $p_3$. Thus, it becomes conceptually possible that there exists an intervention $I_M = i_M$ such that it changes $M$'s value while it does not change $P$'s value. This, in turn, means that it also becomes possible to consistently establish that a mental variable $M$ is a cause (simpliciter) of a causal variable $E$, which shows that the interventionist exclusion argument can be blocked.

## 5 Independent motivation and possible objections

Showing that replacing strict by probabilistic supervenience can block the interventionist exclusion argument is one thing. Arguing that this move is reasonable and that assuming probabilistic instead of strict supervenience is well motivated is another. In this section, we provide some independent motivation for probabilistic supervenience and try to dispel possible worries.

### 5.1 Supervenience in the image of causation

We would like to put forward the following independent motivation for probabilistic supervenience. It consists in a brief comparison with causation, historically as well as conceptually. Historically, causation was often viewed as the glue that holds reality together. This view can be traced back to Aristotle (Stein, 2012). Why do some types of events but not others regularly succeed each other in time? How can we understand these patterns though these types of events are not connected conceptually? For example, copper expands when heated, but the concepts of copper and heating do not imply this behavior in any way. The answer is to assume a contingent causal relation that can do that job. Learning about causal relations provides us with information about which types of events do and do not succeed a certain type of event in time and often even with information about the probability with which the effect can be expected to succeed the cause. Vice versa, knowing the underlying causal relation, observing a phenomenon allows us to say which events might and which ones might not have preceded the effect. Thus, causation structures the space of logically possible successions of types of events in both directions, towards the future and the past.

From a formal point of view, supervenience plays a similar role, but vertically rather than horizontally. While causation restricts the space of events that could precede or succeed an event of a certain type, supervenience restricts the space of possible realizers of a state on a lower level as well as the space of states that might emerge from said state (and possibly other states) at a higher level. Like in the case of causation, what realizes what and what emerges from what are not questions that can be answered a priori based on conceptual analysis. These are empirical matters (cf. Leuridan, 2017). Actual research is required, for example, to find out that water is $H_2O$ or to determine which brain processes can constitute which mental states. Summarizing, we can say that while causation anchors or grounds later events in earlier events, supervenience anchors or grounds higher-level states in lower-level states.[15]

If we now take another look at the history of causation and especially at the prevalent view on causation up until the eighteenth century—as held, for example, by Aristotle, Descartes, Spinoza, and Newton—we can observe that causation was typically understood as a strict relation. This relation was similar to what we would now call *metaphysical necessitation*, even though these particular terms were not employed. The idea was that causes fully determine their effects and that this is the only way how causal relations can hold reality together.[16] More recent developments in philosophy, however, showed that there are many causal relations out there that are at best probabilistic. Nowadays there is a whole plethora of probabilistic theories or at least probability-friendly accounts of causation on the market (see, e.g., Cartwright, 1979; Eells, 1991; Pearl, 2000; Spirtes et al., 1993; Suppes, 1970) and many philosophers do not consider it as problematic to accept that causation might be an inherently probabilistic relation.

We propose that it might be time for a similar step forward when it comes to the concept of supervenience. Maybe also strict supervenience is an artefact of a bygone age and it is time now to embrace supervenience in its full richness instead of only focusing on the extreme case. Note that probabilistic supervenience, like probabilistic causation, is still well capable of performing the job outlined above: It still restricts what is possible at higher as well as at lower levels given a certain event or phenomenon.

[15] For a more formal argument for the claim that supervenience shares important features with causation, see (Gebharter, 2017a).

[16] Descartes famously diverged from Aristotle regarding the number of types of causes. He posited only one—the efficient cause—as opposed to Aristotle's four. Despite this difference, both Descartes and Aristotle agreed that causation is a strict relation and that causes necessitate their effects: Once the efficient cause is in place, its effect must follow (Schmaltz, 2008; Stein, 2012). This view was accepted by post-Cartesian philosophers and scientists until Hume, who was the first to question the received view. Whether Hume denied that the relation of causality itself is a metaphysical necessity relation, or whether he only denied that we can ever be justified in making particular causal claims (which, according to him, were indeed claims about the necessitation relation between the cause and the effect), is a matter of dispute (cf. Beebee, 2011). However, Hume started the progressive weakening of the modality of causal relations, paving the way for today's probabilistic approaches.

Downwards looking, even strict supervenience allows for more than one lower-level realizer of a given phenomenon due to multiple realizability. Maybe it is time to allow for the same one-to-many relation when looking upwards as well. This could still constrain the space of logically possible higher-level states that might emerge from a given lower-level state, the only difference being that it would in principle allow for more than one such state.

### 5.2 Probabilistic supervenience vs. causal closure

In this subsection we want to discuss a possible conceptual worry one might have about probabilistic supervenience. Assume that we find a case in which it turns out that, according to interventionism, a mental variable $M$ is causally efficacious w.r.t. a causal variable $E$. Now one might worry that this finding would stand in contradiction to another typical naturalistic assumption: the closure of the physical domain. Physical closure says that any physical state can be sufficiently explained by pointing to its physical causes. Citing a non-physical cause—that would be $M$ in our example—-cannot add anything to whether such a state is instantiated. For the moment, let $E$ model the physical state to be explained. In terms of interventionism this would mean that once the value of $P$—let us assume that $P$ is $E$'s sufficient physical cause—is fixed, intervening on $M$ cannot have any influence on $E$ anymore. This is true if we assume strict supervenience since, as we saw earlier, $M$'s value cannot be changed once $P$'s value is fixed. However, the same does not hold if we assume probabilistic supervenience. Assuming the latter allows in principle for an intervention on $M$ w.r.t. $E$ that is associated with a change of $E$'s value even if $P$'s value does not change. But then, obviously, $M$ can have a causal influence on $E$ over and above its sufficient physical cause $P$ and, therefore, causal closure would be violated.

We believe that nothing is wrong with the above line of reasoning. Friends of probabilistic supervenience do, however, have several options to respond. The first kind of response they could give is to simply bite the bullet. If the physical domain is causally closed, then the mental cannot add anything and vice versa: If the mental causally contributes to what is going on at the physical level, then causal closure cannot be upheld. As a metaphysician, one needs to make a choice here. Either go for causal closure and throw mental to physical causation out the window, or the other way round. It is not that unusual that supporters of mental causation decide in favor of the latter option, may that be more explicitly or rather implicitly. Take List and Menzies' (2009) argument in favor of mental causation as an example. In their account, an event's mental causes can even exclude its physical causes altogether.

One could also give a somewhat more nuanced answer when deciding in favor of the closure of the physical domain. One might drop mental to physical causation but keep mental causation in the form of mental to mental causation. In terms of interventionism, one could formulate this as follows: Because the physical domain is causally closed, intervening on $M$ will not lead to any change in $E$ if $E$ stands for a physical state. However, if $E$ represents a mental state, then intervening on $M$ might well lead to a change in $E$ even if $M$'s supervenience base $P$ does not change. Nothing we assumed so far excludes that such a change in $E$ associated with an intervention on $M$ must change $E$'s physical supervenience base $P^*$ as well. But if this is possible and we can change the mental variable $E$'s value by means of intervening on $M$ without adding anything to $P$'s causal effect on $P^*$, then physical closure is not violated at all. Keep in mind that if we go for this route, the mental is still sufficiently anchored or grounded in the physical. It still does not follow that anything goes at the mental level; the mental variable $E$'s values can still be sufficiently constrained by $E$'s supervenience base $P^*$.

### 5.3 Probabilistic supervenience and physicalism

Finally, let us discuss another possible worry one might have about probabilistic supervenience.[17] Typically, causal exclusion arguments are launched based on the core assumptions of non-reductive physicalism and are aimed at showing that non-reductive physicalism is untenable and should be discarded in favor of reductive physicalism. Now one might worry that replacing strict supervenience by probabilistic supervenience renders one's position a non-physicalist position right from the beginning. Thus, countering the interventionist causal exclusion argument by relying on probabilistic supervenience would not help one since it would at the same time mean to give up the position one wanted to defend, viz. non-reductive physicalism.

To counter this worry, let us first remind the reader of what we announced in the introduction and saw in full detail in Sect. 3: The interventionist causal exclusion argument does not require all the premises classical exclusion arguments require. In particular, it only requires the two assumptions that the mental is non-identical to the physical and that the mental (strictly) supervenes on the physical. While these assumptions are both crucial theses of non-reductive physicalism, they are not sufficient for a position to be a type of non-reductive physicalism. Non-reductive physicalism does typically also crucially depend on the thesis that the physical domain is causally closed, that every causal event has a causally sufficient physical cause or explanation. We already discussed how probabilistic supervenience relates to the causal closure of the physical domain in sub Sect. 5.2. However, since the interventionist exclusion argument does not require that latter assumption, it does not only threaten non-reductive physicalism, but any philosophical view committed to mental causation that relies on the non-identity of the mental and the physical and the assumption that the mental supervenes on the physical, including some dualist positions. Thus, a way to block the interventionist exclusion argument is not only attractive for non-reductive physicalists.

Be that as it may, one might still worry how a subscriber to non-reductive physicalism can consistently uphold their position after exchanging strict supervenience for its probabilistic counterpart. One specific worry may be that probabilistic supervenience leads to violations of the causal closure of the physical domain. As we saw in sub Sect. 5.2, probabilistic supervenience does not automatically imply this. It is conceptually compatible with all of the following possibilities: (i) Mental variables causally influence only other mental variables, (ii) mental variables causally influence only physical variables, (iii) mental variables sometimes causally influence both other mental and physical variables, and (iv) mental variables causally influence no

[17] We are indebted to an anonymous reviewer for pushing us to reflect more on this possible worry.

other variable whatsoever. Which of these possibilities is actually the case is a contingent matter that we cannot decide simply by endorsing probabilistic supervenience. It depends on the world, in particular on whether the interventions supporting each of these possibilities exist which, in turn, depends on the specific conditional probabilities of the mental variables' values given their physical supervenience base variables' values to be found in the world. Now possibility (i) does not violate causal closure. Thus, if the world is such that (i) comes out as true, one can clearly uphold non-reductive physicalism. Since probabilistic supervenience does not determine which one of these four possibilities comes out as true, replacing strict supervenience by probabilistic supervenience does not automatically render one's position different from non-reductive physicalism.

Summarizing what we found so far, non-reductive physicalism is in principle compatible with probabilistic supervenience. There is still concern that the claim that several mental states are now compatible with a single physical state automatically makes one a dualist. This would go against another crucial requirement of non-reductive physicalism: Non-reductive physicalism is typically understood as a monist position. A strategy to dispel such a worry goes as follows: One way to conceptualize physicalism is to assume that all the entities in the world are physical. Thus, physicalism is an entity monism. To acknowledge another crucial commitment of non-reductive physicalism, viz. that the mental is not identical to the physical, one can then assume that some properties are mental while others are physical, but that both are instantiated by physical entities. Thus, one ends up with a position that is monist when it comes to entities (they are all physical), but dualist when it comes to properties (they are either mental or physical). Hence, one's version of non-reductive physicalism is a monism in a relevant sense: There are only physical entities. At the same time, it is non-reductive because probabilistic supervenience provides a metaphysical foundation for another intuition crucial to non-reductive physicalists: The intuition that at least some mental properties, such as phenomenal experiences, cannot be reductively explained by any physical properties (Kim, 2005; Levine, 1983). If the relation between mental and physical properties is probabilistic rather than strict supervenience, it becomes clearer why certain mental properties, while supervening on physical properties, cannot be reductively explained: There exists a metaphysical possibility for slightly different mental properties to supervene on the same physical base.

Finally, we acknowledge that one might still have worries. For example, not everyone believes that even strict supervenience is a strong enough assumption for physicalism. Kroedel and Schulz (2016), for example, propose that grounding should be used in causal exclusion contexts. If one considers strict supervenience as too weak, then surely probabilistic supervenience will be even worse. These are interesting and original objections, but since they do not specifically threaten probabilistic supervenience, but supervenience in general, we believe that this paper is not the right place to discuss them in more detail. For now, we content ourselves with emphasizing that the majority view still relies on supervenience as an essential assumption for non-reductive physicalism and with pointing to our argumentation at the end of Sect. 2 for why probabilistic supervenience still plays the same role as strict supervenience for anchoring or grounding the mental in the physical.

## 6 Conclusion

We argued that previous interventionist approaches to saving mental causation from the causal exclusion argument do not work because of the assumption of strict supervenience. To this end, we put forward what we believe to be the strongest version of the interventionist causal exclusion argument currently on the market. From the premises required for the classical exclusion argument it only requires the assumption that the mental is different from and non-trivially supervenes on the physical. In addition, it works with the weaker notion of an intervention variable **(IV)*** and even under the assumption that all variable sets used for evaluating causal claims need to satisfy independent fixability **(IF)**.

As we saw, strict supervenience renders any intervention on any mental variable impossible. Consequently, no mental variable can be a cause of any other variable. However, if the assumption of strict supervenience is replaced by the assumption of probabilistic supervenience, the values of mental variables are not fully determined by the values of the variables representing their physical supervenience bases, which renders interventions on mental variables possible. Consequently, if probabilistic supervenience is accepted, then the mental can be causally efficacious within the interventionist account of causation. At this point we would like to emphasize that this does not yet establish that there actually is any mental causation. It only paves the road for mental causation by showing that it is compatible with the other assumptions used in the interventionist causal exclusion argument.

Moreover, probabilistic supervenience would not only save mental causation from the causal exclusion argument, but also has other virtues. First, it is compatible with the naturalistic worldview because it shows that mental properties are not completely disconnected from their underlying physical properties but are probabilistically anchored in them. Second, it allows for symmetric multiple realizability: Not only the same mental state can be realized by different physical states, but also different mental states can be realized by the same physical state. This provides a metaphysical foundation for qualitatively different subjective mental states given the same underlying physical states, and thereby grounds the intuition behind the explanatory gap argument (Levine, 1983). Finally, it leads to new conceptual developments in the debate on mental causation, because it shows that the thesis of the causal closure of the physical is compatible with the thesis of the causal efficacy of the mental only if the causal efficacy of the mental is restricted to the domain of the mental. These considerations do not conclusively show that the mental actually probabilistically supervenes on the physical. What we wanted to achieve is rather to offer a metaphysically possible and somewhat plausible in-principle way to block the interventionist causal exclusion argument. We argued that probabilistic supervenience is a possibly useful concept for supporters of mental causation and that it is not that outlandish as it might seem at first glance.

**Acknowledgements** We would like to thank two anonymous reviewers for helpful comments on an earlier version of this paper. This work was funded by the German Research Foundation (DFG), project number 441311834 (Alexander Gebharter) and Research group Inductive Metaphysics (FOR 2495), research grant SCHU 1566/11-2 (Maria Sekatskaya).

**Funding** Open access funding provided by Università Politecnica delle Marche within the CRUI-CARE Agreement.

## Declarations

**Conflict of interest** The authors declare that there are no conflicts of interest.

## References

Baumgartner, M. (2010). Interventionism and epiphenomenalism. *Canadian Journal of Philosophy, 40*(3), 359–383.

Baumgartner, M. (2013). Rendering interventionism and non-reductive physicalism compatible. *Dialectica, 67*(1), 1–27.

Baumgartner, M. (2018). The inherent empirical underdetermination of mental causation. *Australasian Journal of Philosophy, 96*(2), 335–350.

Baumgartner, M., & Drouet, I. (2013). Identifying intervention variables. *European Journal for Philosophy of Science, 3*(2), 183–205.

Baumgartner, M., & Gebharter, A. (2016). Constitutive relevance, mutual manipulability, and fat-handedness. *British Journal for the Philosophy of Science, 67*(3), 731–756.

Beebee, H. (2011). *Hume on Causation*. Routledge.

Cartwright, N. (1979). Causal laws and effective strategies. *Nous, 13*, 419–437.

Craver, C. (2017). Stochastic supervenience. In M. P. Adams, Z. Biener, U. Feest, & J. A. Sullivan (Eds.), *Eppur si muove: Doing history and philosophy of science with Peter Machamer: A collection of essays in honor of Peter Machamer* (pp. 163–170). Springer.

de Grefte, J., & Gebharter, A. (2021). The causal theory of knowledge revisited: An interventionist approach. *Ratio, 34*(3), 193–202.

Eells, E. (1991). *Probabilistic Causality*. Cambridge University Press.

Eronen, M., & Brooks, D. (2014). Interventionism and supervenience: A new problem and provisional solution. *International Studies in the Philosophy of Science, 28*(2), 185–202.

Gebharter, A. (2017a). Causal exclusion and causal Bayes nets. *Philosophy and Phenomenological Research, 95*(2), 353–375.

Gebharter, A. (2017b). Causal exclusion without physical completeness and no overdetermination. *Abstracta—Linguagem Mente E Ação, 10*, 3–14.

Gebharter, A., Graemer, D., & Scheffels, F. H. (2019). Establishing backward causation on empirical grounds: An interventionist approach. *Thought: A Journal of Philosophy, 8*(2), 129–138.

Hitchcock, C. (2012). Theories of causation and the causal exclusion argument. *Journal of Consciousness Studies, 19*(5–6), 40–56.

Hoffmann-Kolss, V. (2014). Interventionism and higher-level causation. *International Studies in the Philosophy of Science, 28*(1), 49–64.

Hoffmann-Kolss, V. (2022). Interventionism and non-causal dependence relations: New work for a theory of supervenience. *Australasian Journal of Philosophy, 100*(4), 679–694.

Kim, J. (1973). Causes and counterfactuals. *The Journal of Philosophy, 70*, 570–572.

Kim, J. (2005). *Physicalism, or something near enough*. Princeton University Press.

Kinney, D. (2023). Bayesian networks and causal ecumenism. *Erkenntnis, 88*(1), 147–172.

Kistler, M. (2017). Higher-level, downward and specific causation. In M. P. Paoletti & F. Orilia (Eds.), *Philosophical and scientific perspectives on downward causation.* Routledge.

Kroedel, T., & Schulz, M. (2016). Grounding mental causation. *Synthese, 193*(6), 1909–1923.

Leuridan, B. (2012). Three problems for the mutual manipulability account of constitutive relevance in mechanisms. *British Journal for the Philosophy of Science, 63*(2), 399–427.

Leuridan, B. (2017). Supervenience: Its logic and its role in classical genetics. *Logique Et Analyse, 198*, 147–171.

Leuridan, B., & Lodewyckx, T. (2020). Diachronic causal constitutive relations. *Synthese, 198*, 9035–9065.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly, 64*, 354–361.

Lewis, D. (1986). *Philosophical Papers: Volume II*. Oxford University Press.

List, C., & Menzies, P. (2009). Non-reductive physicalism and the limits of the exclusion principle. *Journal of Philosophy, 106*(9), 475–502.

McLaughlin, B., & Bennett, K. (2021). Supervenience. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy.* Stanford: Metaphysics Research Lab. Retrieved from https://plato.stanford.edu/archives/sum2021/entries/supervenience/

Pearl, J. (2000). *Causality*. Cambridge University Press.

Schmaltz, T. (2008). *Descartes on causation*. Oxford University Press.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, Prediction, and Search*. Springer.

Stein, N. (2012). Causal necessity in Aristotle. *British Journal for the History of Philosophy, 20*(5), 855–879.

Stern, R., & Eva, B. (2023). Anti-reductionist interventionism. *British Journal for the Philosophy of Science, 74*(1), 241–267.

Suppes, P. (1970). *A Probabilistic Theory of Causality*. North-Holland.

Woodward, J. F. (2003). *Making Things Happen*. Oxford University Press.

Woodward, J. F. (2008). Response to Strevens. *Philosophy and Phenomenological Research, 77*(1), 193–212.

Woodward, J. F. (2015). Interventionism and causal exclusion. *Philosophy and Phenomenological Research, 91*(2), 303–347.

Woodward, J. F. (2016). The problem of variable choice. *Synthese, 193*(4), 1047–1072.