



Closure, counterfactualist causation, and Zhong's new causal argument for physicalism

Jessica Wilson¹ 

Received: 17 May 2025 / Accepted: 25 July 2025

© The Author(s), under exclusive licence to Springer Nature B.V. 2025

Abstract

I assess Zhong's new "causal argument" for physicalism, which differs from previous such arguments in that the premises and conclusion pertain (not just to physical, but) to "physically acceptable" entities or features, which may be either physical or "grounded by" (i.e., metaphysically dependent on) the physical. Zhong argues that his new causal argument improves on previous versions in that the conclusion (unlike previous causal arguments, he maintains) supports non-reductive as well as reductive versions of physicalism, and in that the premises of his argument are better motivated than those of the original arguments. I argue that neither of these motivations are in place. Along the way, I offer a new reason to reject non-contrastive counterfactual accounts of causation.

Keywords Physicalism · Causal closure · Counterfactual accounts of causation · Causal arguments for physicalism

Overview

In his paper (2023), Zhong offers a new "causal argument" for physicalism—about mental properties, in particular:

Physicalism (about the mind) is the view that mental properties are metaphysically dependent on physical properties—in other words, the mental is either identical to or grounded by the physical. Here, the term "physical properties" is used in the standard, narrow sense, referring to properties posited by ideal physics (as well as their aggregations and configurations). [...] Why should we accept mental–physical metaphysical dependence rather than a weaker kind of mental–physical correlation (e.g., merely nomological dependence)? In this article, I shall develop a causal argument for physicalism, which aims to establish

✉ Jessica Wilson
jessica.m.wilson@utoronto.ca

¹ Department of Philosophy, University of Toronto, Toronto, ON, Canada

that only physicalism can accommodate the causal efficacy of mental properties.

(1)

As Zhong observes, a causal argument for physicalism is already on the books—notably, as presented or discussed by Kim (1998), Papineau (2001), and others—but as Zhong sees it, the version of physicalism targeted in this previous argument is *reductive* physicalism, according to which every mental feature (property, state, etc.) is identical to some or other physical feature, as per the following reconstruction:

- 1a *Interaction*: Mental properties are causally relevant to some physical effects.
- 1b *Closure*: Every physical effect has a sufficient physical cause (at any time at which it has a cause).
- 1c *Exclusion*: There is no systematic overdetermination by physical and non-physical causes.
- ∴ *Identity*: Therefore, mental properties are physical properties.

Call this “the original causal argument.” As Zhong sees it, the conclusion of the original causal argument “threatens non-reductive physicalism” and more generally “cannot help explain why some anti-reductionists accept the general doctrine of physicalism” (3). Moreover, Zhong sees the premises of the original causal argument as subject to challenge. And he suggests that a better causal argument for physicalism—better both in that its conclusion is compatible with non-reductive as well as reductive versions of physicalism, and in that its premises are more defensible—can be given, as follows:

- 2a *Interaction**: Mental properties are causally relevant to some physically acceptable effects.
- 2b *Closure**: Every physically acceptable effect has a sufficient physically acceptable cause (at any time at which it has a cause).
- 2c *Exclusion**: There is no systematic overdetermination by physically acceptable and physically unacceptable causes.
- ∴ *Identity**: Therefore, mental properties are physically acceptable properties.

Call this “the new causal argument.” Here, the key difference with the original causal argument is that the premises and conclusion pertain (not just to physical, but) to “physically acceptable” entities or features, where to be physically acceptable is to be either physical or “grounded by” (i.e., completely metaphysically dependent on) the physical. The remainder of Zhong’s paper is devoted to arguing that each of the premises of the new causal argument is defensible, and more so than those of the original causal argument.

Zhong’s intriguing suggestion raises several important and interesting issues pertaining to the support for and formulation of physicalism. My remarks here focus on his stated motivations for the new causal argument. I’ll first argue (§1) against Zhong’s claim that previous causal arguments (including the original causal argument) are arguments for distinctively reductive physicalism. I’ll next argue (§2) against Zhong’s reasons for thinking that *Closure** is to be preferred over *Closure*, the most important of which is that *Closure**, but not *Closure*, is compatible with a counterfactual

account of causation, and the associated possibility that some physical effects have non-physical but physically acceptable causes. Here I'll observe that contrastive versions of such an account are no threat to *Closure*, and that in any case counterfactual accounts (including the non-contrastive interventionist variation that Zhong favors) face a plethora of well-known difficulties. I'll moreover offer a new objection to a non-contrastive counterfactual account, according to which such an account problematically breaks the systematic connection between causes and laws. I'll finally argue (§3) that Zhong is too hasty in claiming that physicalists can happily accept *Closure** and the associated possibility of some physical effects' having only non-physical causes "if the non-physical causes are physically grounded." For as I'll argue, the only working strategies for establishing such grounding/metaphysical dependence claims proceed by establishing that the non-physical goings-on at issue do not have any causal powers beyond those had by their physical dependence base. I conclude that physicalists do better to stick with the causal argument(s) already on the books.

1 Which version of physicalism does the original causal argument support?

My first remarks concern Zhong's claim that the original causal argument (as presented above) is aimed at establishing reductive physicalism, which claim is supposed to provide strong motivation for the new causal argument, understood as explaining "why some anti-reductionists accept the general doctrine of physicalism" (71). As I'll now argue, the original causal argument is better seen as supporting non-reductive physicalism, and has been offered as such by several philosophers. Moreover, I'll argue, a version of this argument more in line with that offered by Kim and Papineau is best seen as an argument not for reductive physicalism, but for physicalism in general, and again is offered as such by several philosophers (including Papineau).

To proceed, it is worth first revisiting the sort of causal argument offered by Kim:

Suppose that a certain [mental] event [...] causes a physical event. The causal closure of the physical domain says that this physical event must also have a physical cause. [...] What is the relationship between the two causes, one mental and the other physical? [It might be suggested that] the mental cause and the physical cause are each only a partial cause, and that they together make up one sufficient cause. This surely is an absurd thing to say, and in any case, it violates the causal closure principle. [...] Could it be that the mental cause and the physical cause are each an independent sufficient cause of the physical effect? The suggestion is that the physical effect is overdetermined. So if the physical cause hadn't occurred, the mental cause by itself would have caused the effect. This picture is again absurd [given the synchronic dependence of the mental on the physical, and] violates the causal closure principle as well: in the counterfactual situation in which the physical cause does not occur, the closure principle is violated [...]. These two ways of looking at the situation are obvious non-starters. [...] Given that any physical event has a physical cause, how is a mental cause also possible? (Kim 1989: 280–81)

Kim clearly and repeatedly maintains that the best way out here is to identify the mental and physical features. So there is at least one causal argument on the books which at least some (notably, Kim) have seen as establishing a specifically reductive version of physicalism.

Note, however, that Kim's causal argument differs from the original causal argument, in that the "interaction" premise in Kim's argument simply registers that a mental feature *M* *causes* a physical effect, whereas the "interaction" premise in the original causal argument registers that mental feature *M* is *causally relevant* to a physical effect, where (as Zhong specifies in note 2) "[w]hen an event *c* that instantiates a property *F* causes an effect *e* in virtue of having *F* [...] *F* is causally relevant to *e*." As I'll now discuss, this difference makes a difference to what conclusion, and associated version of physicalism, is best seen as supported by the original causal argument.

Indeed, Kim's argument is best seen as raising to salience the question of how a mental feature can be efficacious *at all vis-à-vis* a physical effect, given that, by *Closure*, any such effect already has a physical cause, and given the implausibility of overdetermination. Kim's preferred answer offers a straightforward means of gaining such efficacy—namely, by identifying *M* with *P*, as per the usual reductive physicalist response. But as Horgan (1989) observes, this response leaves open what is arguably the deeper question underlying the problem of mental causation: "even if individual mental events and states are causally efficacious, are they efficacious *qua* mental?" (47). As Shoemaker (2001) expresses the concern:

If what is in fact an instance of a mental property causes something ... but does so in virtue of being an instance of a physical property rather than in virtue of being an instance of a mental property, then the causal efficacy of the mental does not seem to have been adequately vindicated. (28)

These considerations indicate that proper accommodation of the efficacy of the mental requires that mental properties be efficacious *qua* mental—in other words, that a mental property be efficacious *in virtue of* being mental, and not just efficacious in virtue of being identical to some efficacious physical property.

Now return to the original causal argument, and the associated *Interaction* premise, according to which some mental property *M* is causally relevant to some physical effect *P**. Taking Horgan's point and Zhong's definition of causal relevance into account, *Interaction* is naturally understood as expressing that *M* causes *P** in virtue of being *M*—that is, *qua* mental, and where room for such mental quausation (per Horgan's evocative expression) involves, at a minimum, *M*'s being distinct from its cotemporal dependence base feature *P*. Perhaps *Interaction* doesn't force such a "qua" reading; but my present point is just that in the existing literature, endorsement of mental causal relevance is commonly understood as pushing against reductive physicalism, and moreover (as in Horgan's and Shoemaker's discussions, among others) as pushing, via one or other strategy for gaining mental causal relevance, for non-reductive physicalism.

But can non-reductive physicalists accept the other premises of the original causal argument? They can, and often do. To start, as physicalists, they standardly accept *Closure*—hence in the cases at hand, they grant that, notwithstanding that *M* causes

P^* , P also causes P^* . But how can they accept this, compatible with *Exclusion*? Here it is crucial to observe that, as in Kim's statement of his causal argument, the support for *Exclusion* reflects that mental causes are not plausibly taken to systematically overdetermine effects already brought about by lower-level physical goings-on *in the same way as firing squad or double-rock-throw cases*. Only if one is thinking of overdetermination along double-rock-throw lines would overdetermination entail—what Kim described as “absurd,” given the synchronic dependence of the mental on the physical—that “if the physical cause hadn't occurred, the mental cause *by itself* would have caused the effect” (*italics inserted*). Correspondingly, *Exclusion* should be understood as expressing the more specific claim that “There is no systematic double-rock-throw-style overdetermination by physical and non-physical causes.”

Non-reductive physicalists of the “causal compatibilist” variety can happily accept both *Closure* and *Exclusion*, so understood; and indeed a common strategy of non-reductive physicalists—one nicely dovetailing with their strategy for achieving mental causal relevance—involves advancing a realization relation between mental properties and their physical dependence base properties which is sufficiently intimate that even if mental property M and physical property P each end up being a cause of a single physical effect P^* , the resulting overdetermination is of a benign, non-double-rock-throw variety, compatible with *Exclusion*. For example, on the proper-subset-of-powers approach to realization (Wilson, 1999; Shoemaker, 2001; Clapp, 2001), a mental property has, on any given occasion, a proper subset of the token powers had by the physical property upon which it coterminally depends on that occasion. On this view, when a mental property causes a physical effect also caused by its realizer, there is only one power exercised (in contrast with double-rock-throw cases), compatible with both the mental property and the physical property being causes of the effect (reflecting that each has and exercises the power in question). On this approach, though only one power is exercised in cases of mental causation, it nonetheless remains that the mental property is efficacious *qua* mental, in having a distinctive power profile, tracking difference-making considerations or systems of laws which abstract away from various microphysical details (including the more specific powers had by the mental property's physical realizer). And as I have argued (see, e.g., Wilson, 2021), many specific non-reductive physicalist accounts—e.g., those appealing to functional roles, or the determinable/determinate relation—are properly seen as implementations of a subset-of-powers approach to realization, and the associated response to Kim-style exclusion concerns.

Of course, whether this or other variations on the non-reductive physicalist theme are successful is a further issue.¹ But my present point is just that it is not the case, as Zhong claims, that the original causal argument “threatens non-reductive physicalism” and more generally “cannot help explain why some anti-reductionists accept the general doctrine of physicalism” (71). On the contrary, the premises of this argument are commonly understood as supporting non-reductive rather than reductive versions of physicalism.

¹ See Wilson (2021) for discussion of and responses to potential objections to a proper-subset-of-powers approach to realization, under the rubric of “Weak emergence.”

Could Zhong's dialectical narrative, according to which previous causal arguments are arguments for reductive rather than non-reductive physicalism, be revived by revising *Interaction* in his framing of the original causal argument to express (not that M is causally relevant to physical effect P^* , but rather) just that M causes P^* (as in Kim's causal argument)? The answer is "no." For once it is appreciated that there is more than one way to satisfy *Exclusion*—either by way of identity or by way of non-reductive realization—Kim's causal argument can be seen as compatible with either a non-reductive physicalist or a reductive physicalist conclusion. Kim goes for identity and hence for reductive physicalism, but that isn't forced by the premises of Kim's causal argument. Hence, Kim's causal argument, properly understood, is *already* an argument for the more general conclusion of Zhong's new causal argument.

That this is so is supported by attention to, e.g., Papineau's (2001) discussion. There, Papineau starts by registering that physicalism comes in non-reductive as well as reductive varieties, and that his aim is to "understand the reasons for physicalism of any kind":

[C]ontemporary physicalism is an ontological rather than a methodological doctrine. It claims that everything is physically constituted, not that everything should be studied by the methods used in physical science. [...] This methodological liberalism goes with the fact that the ontological claims of fin-de-siècle physicalism are often carefully nuanced. If physicalism simply meant type-type physical reduction [...] then methodological unity of science would arguably follow, in principle at least, from physicalism. But physicalism today clothes itself in various subtler shades. We have physical supervenience, physical realization, token-token physical identity, and so on. [...]

My concern here is not to distinguish the different species of physicalism [...] but to try to understand the reasons for physicalism of any kind. Why have so many analytic philosophers in the second half of the twentieth century suddenly become persuaded that everything is physical? (4)

He then goes on to present Kim's causal argument as the core argument motivating any of these "species of physicalism."

Summing up: insofar as the *Interaction* premise of (what Zhong presents as) the original causal argument requires that M (not just cause but moreover) be causally relevant to P^* , it is an argument for non-reductive, not reductive, physicalism. And if the *Interaction* premise in the original causal argument is revised to require only that M cause P^* , as in Kim's causal argument, the result is (as in Papineau's discussion) an argument for "physicalism of any kind" or "any species of physicalism." None of this is to say that Zhong's new causal argument isn't new; it certainly is. But contra Zhong's claim, we don't *need* his argument in order to make room for non-reductive physicalism—either the original causal argument or Kim's causal argument can do that trick.

2 Is *Closure* undercut by a counterfactual account of causation?

It remains that Zhong's new causal argument might be an improvement, one way or another, on previous causal arguments for physicalism in general, or non-reductive physicalism in particular. Indeed, Zhong suggests that each premise of his new causal argument is more plausible than the parallel premise in the original causal argument.² Here I want to focus on Zhong's replacement of

1b *Closure*: Every physical effect has a sufficient physical cause (at any time at which it has a cause).

with

2b *Closure**: Every physically acceptable effect has a sufficient physically acceptable cause (at any time at which it has a cause).

Zhong maintains that *Closure** is more plausible than *Closure*, on grounds, first, that the empirical motivation for *Closure* depends on acceptance of an unsatisfactory physical (transference or conserved quantity-based) account of causation, and second, that by lights of the sort of counterfactual account of causation that Zhong endorses, *Closure* is likely false.

In what follows, I'll argue that these cases for *Closure**'s being more plausible than *Closure* are unconvincing.

Let's start with Zhong's claim that empirical support for *Closure* relies on accepting an unsatisfactory account of causation:

It has been assumed that the laws of conservation of energy and momentum, a cornerstone of contemporary physics, provides strong empirical support to the causal closure of physics (Papineau, 2001; Gibb, 2010). But obviously, physical conservation as such does not entail physical closure. To move from the conservation of the physical domain to the causal closure of physics, one needs to presuppose a physical account of causation, such as the conserved quantity theory of causation, according to which causal facts are reducible to facts about the processes of transferring conserved physical quantities (Salmon, 1984; Dowe, 2000). However, the physical account of causation is highly controversial in that it fails to distinguish between causally relevant and irrelevant parts of a physical process. [...] Moreover, a physical theory of causation has difficulty explaining how we can have causal knowledge. We normally form our causal beliefs without knowing the underlying physical processes, let alone the processes of transferring conserved physical quantities. (6)

Transference and other physical accounts of causation do face various difficulties. Be this as it may, in any case the empirical support for *Closure* does not depend on endorsing such an account of causation.

Here it is useful to attend to Papineau's very thorough (2001) discussion of *Closure* and its motivations. To be sure, Papineau's "original thought" was that *Closure* was based in conservation laws:

² Or, presumably, than the parallel premises in Kim's causal argument.

My original thought was that the completeness of physics would follow from the fact that physics can be formulated in terms of conservation laws. If the laws of mechanics tell us that important physical quantities are conserved regardless of what happens, then doesn't it follow that the later states of physical systems are always fully determined by their earlier physical states? (14)

But Papineau goes on to argue that the empirical motivation for *Closure* does not rest on physics's involving conservation laws, since, e.g., "the doctrine of the conservation of energy is itself consistent with the existence of special [e.g., mental] forces, as long as those forces are conservative" (28).

What Papineau ends up concluding is that it is not conservation laws, but rather attention to why physicists came to posit conservation laws, which points towards one of two empirical motivations for *Closure*:

[M]y thought here is that the arguments behind the conservation of energy give inductive reason to suppose that all forces reduce to a small number of fundamental forces. We have already seen how Helmholtz's formulation of the conservation of energy hinged on the assumption that friction and other dissipative forces are nonfundamental forces, macroscopic manifestations of processes involving more fundamental conservative forces. For it is only if we see macroscopic forces such as friction as reducing to fundamental conservative forces that we can uphold the universal conservation of energy. Now, this point can be viewed as providing inductive support for the general thesis that all apparently special forces reduce to a small stock of fundamental forces. The special forces that have been quantitatively analyzed, such as friction, turn out to reduce to more fundamental conservative forces. So this provides inductive reason to conclude that any other apparently special forces, such as muscular forces, vital forces, or mental forces, will similarly reduce. (28)

The connection between the reductions of macroscopic forces to fundamental physical forces and *Closure* is made explicit in what Papineau calls "the argument from fundamental forces" for that thesis:

The first argument is that all apparently special forces characteristically reduce to a small stock of basic physical forces that conserve energy. [...] So, while we ordinarily attribute certain physical effects to "muscular forces," say, or indeed to "mental causes," we should recognize that these causes, just as all causes of physical effects, are ultimately composed of the few basic physical forces. (27)

That all physical effects are ultimately caused by fundamental physical forces supports "the completeness of physics, by which I mean that all physical effects are due to physical causes" (27)—that is, *Closure*.

Papineau also observes a second motivation for *Closure*, which he calls "the argument from physiology":

The earlier argument suggested that at least most natural phenomena, if not all, can be explained by a few fundamental physical forces. This focused on the

issue of what kind of evidence would demonstrate the existence of extra mental or vital forces. For once we know that other forces exist, then we will know which anomalous accelerations would indicate the presence of special mental or vital forces. Against this background, the argument from physiology is then simply that detailed modern research has failed to uncover any such anomalous physical processes. (30–1)

Papineau's diagnosis of the broadly empirical cases for *Closure* has been widely accepted. For example, in a recent discussion of the mind-body problem, Balog (2025) registers both the argument from fundamental forces and the argument from physiology as the primary motivations for *Closure* and the associated commitment to one or other form of physicalism (a.k.a., "materialism"):

Some contemporary philosophers hold on to Cartesian dualism by denying the Causal Closure Principle. But this has become an increasingly difficult position to maintain, especially after the scientific progress in physiology made by the middle of the 20th century. Whereas Newton allowed forces other than impact and, in principle, as far as physics was concerned, there was no barrier to the existence of special forces such as chemical, biological, or mental, it became clear by the middle of the 20th century that none of those forces are *sui generis*, that is, they are all explicable in terms of the forces we are familiar with from physics. During this time, in addition to discoveries in biochemistry, neurophysiological research mapped the body's neuronal network and analyzed the electrical mechanisms responsible for neuronal activity. If there were *sui generis* mental forces operating inside living bodies, they should have shown up in some way in our observations of neural activity. But research has failed to uncover evidence of anything except familiar physical forces. By the middle of the last century, belief in an immaterial soul that is a *sui generis* cause of action had steadily declined, and materialism had become the dominant view in philosophy and increasingly among the wider public.

Now, I want to make three observations about these two broadly empirical motivations for *Closure*. First, neither appeals to a physical (transference or conserved quantity) account of causation, or to anything that presupposes the truth of such an account. To the extent that attention to conserved quantities or associated conservation laws enters into the motivations for *Closure*, this is due not to commitment to a specific (much less a specifically physical) account of causation, but rather to an inductive generalization from cases wherein seemingly macroscopic forces are revealed to be conserved, because reducible to fundamental physical forces. Second and relatedly, if there is a salient supposition at issue in these motivations for *Closure*, it is that the causation at issue involves the operation of forces or interactions, such that if some forces or interactions entering into the production of a physical effect turn out to be ultimately reducible to fundamental physical forces or interactions, then the latter (or associated entities or features) are plausibly seen as causing the physical effect in question.³ That

³ Note that the line of here is so far neutral on whether macroscopic forces (or associated entities or features) can be seen as also causing the physical effect.

at least some cases of causation involve forces or interactions is compatible with versions of all the main accounts of causation, including, e.g., non-Humean accounts on which causation involves mutual manifestation of dispositions, Humean accounts on which causation is a matter of appropriate regularities, and—a point that will be relevant down the line—counterfactual accounts on which causes are difference-makers. Third, although from a general physicalist perspective, what it is most crucial to avoid are *sui generis* (i.e., fundamental) mental forces, it is worth observing that each of the empirical motivations for *Closure* presupposes and inductively builds on empirical results according to which the physical effects of physically acceptable (e.g., “mechanical” and “muscular”) causes have physical (as opposed to non-physical but physically acceptable) causes. Consequently, on Papineau’s (and Balog’s) construal, these empirical motivations are specifically motivations for *Closure*, not *Closure**.

I turn now to Zhong’s second plank in his case against *Closure**: namely, that by lights of a counterfactual account of causation, *Closure* might be false:

[I]f we adopt a counterfactual theory of causation, it remains open whether every physical effect counterfactually depends on an earlier physical state. It could be the case that a physical effect sometimes counterfactually depends on a higher-level (e.g., mental and social) phenomenon but not on its physical realizer. In such a situation, the mental downwardly causes the physical effect (in a non-overdetermining way)—*Closure* is false (Zhong, 2019). (7)

The possibility of such nonredundant downward causation would be realized whenever the correlation between a cause *c* and an effect *e* is “realization-insensitive”:

To say that a correlation between *c* and *e* is realization-sensitive is to say that if the actual realization base of *c* were to change even slightly, the correlation would no longer hold. In contrast, the correlation between *c* and *e* is realization-insensitive in the sense that the correlation would still hold if the realization base of *c* were to undergo a small change. If the *M–P** correlation is a realization-sensitive case, then the subvenient property *P* is a better candidate than the supervenient property *M* for being the cause of *P**. If it is a realization-insensitive case, then the supervenient property *M* rather than the subvenient property *P* should be regarded as the real cause of *P** (Zhong, 2020). (5)

Plausibly, cases where a multiply realized mental state (say, a feeling of thirst) brings about a physical effect (say, my reaching for the Fresca, or the associated physical basis thereof) are common, in which case this sort of possibility is not only actual but ubiquitous. Since *Closure**, but not *Closure*, makes room for nonredundant non-physical causes of physical effects, *Closure**, Zhong maintains, is correspondingly more plausible.

As regards this line of thought, three points.

First, as it happens, one variety of counterfactual account of causation takes judgments about what counterfactually depends on what to be relative to contrast classes. Such a counterfactual account makes room for the truth of *Closure*, relative to what is plausibly seen as the default contrast class of physical goings-on. As I previously observed:

It is worth noting that even if one endorses a response to Kim-style overdetermination concerns appealing not just to difference-making considerations (as per, e.g., Horgan 1989 and LePore and Loewer 1987 and 1989; see also Bennett 2003) but moreover to a difference-making or counterfactual account of causation (as [per] List and Menzies 2009), one might not be committed to gaining the distinctive efficacy of a higher-level feature *vis-à-vis* a given effect at the cost of denying that its physical dependence base feature is also a cause of the effect (contra Physical Causal Closure); for one may rather suppose that higher-level and lower-level features are each causes, relative to different contrast classes (see, e.g., Hitchcock 1996.) (Wilson 2021, 82)

So in general, acceptance of a counterfactual account of causation does not motivate *Closure** over *Closure*.

Second and in any case, counterfactual accounts of causation are subject to notorious difficulties, associated with cases of preemption and overdetermination (see, e.g., Hall 2004), joint effects (see, e.g., Bennett 1987), probabilistic causation (see, e.g., Menzies 1989), failure to distinguish causal and non-causal dependence (see, e.g., Kim 1973), failure to properly encode the temporal asymmetry of causation (see, e.g., Elga 2000), circularity (see, e.g., Hiddleston 2005), and so on, not to mention controversies concerning the nature and ordering of the unactualized possibilities to which counterfactuals make reference (see, e.g., Fine 1975, Wilson 2014a). Hence, it is that Menzies and Beebe (2024) observe that “[i]ntense discussion over forty years has cast doubt on the adequacy of any simple analysis of singular causation in terms of counterfactuals,” such that “recent years have seen a proliferation of different refinements of the basic idea,” especially in the form of structural equations, causal modeling, or interventionist approaches. Whether these refinements are ultimately any improvement is unclear, however, since they typically remain subject to many of the original difficulties, and moreover introduce new ones.

This is true, in particular, of Woodward’s (2003) interventionist approach, a variation on which Zhong aims to apply to the cases of mental causation at issue here. As Kroedel (2024) describes the approach:

According to interventionism, one event or property causes another if and only if an isolated manipulation of the first event or property would lead to a change in the second event or property while other factors are appropriately held fixed. (2479)

The “holding fixed” requirement is needed to ensure that the causal influence at issue goes through the manipulated variable. But as Kroedel observes, it is unclear how this requirement, hence the interventionist approach, is to be applied to the cases at hand:

What other factors precisely should be held fixed, and at what values? This question becomes especially tricky if we try to apply interventionism to mental causation and at the same time, assume non-reductive physicalism, that is, the view that mental properties are not identical to, but supervene on, physical properties. If we represent both a mental property and the physical property that actually realizes it, should we hold the physical property fixed while interven-

ing on the mental property? The original interventionist theory does not give us an answer because it eschews representing properties that stand in necessary relations, such as the mental property and its physical realizer(s). (2479–80)

Nor, as Baumgartner (2009) argues, can Woodward’s interventionist account be simply extended to the cases at hand by holding fixed the physical realizer while intervening on the mental property: since (all parties agree) the physical realizer necessitates the mental property, such interventions are impossible, rendering the mental property trivially causally inert. As Baumgartner discusses, Woodward’s (2015) attempt to weaken interventionism to allow for non-trivial application to cases of downward mental causation doesn’t succeed, instead rendering “downward causal claims involving supervening variables [...] either meaningless or not analyzable within the interventionist framework” (17). The upshot is that “non-reductive physicalism is in no way supported by interventionism” (17).

Zhong’s preferred strategy of response to this difficulty is, as Kroedel describes it, “nuanced”: roughly, Zhong suggests that to determine whether M is a cause of P^* , we should hold P fixed—not at its actual value (on his binary value model: as *present*), since P ’s being present entails M ’s being present, but rather at P ’s counterfactual value (as *absent*)—and then intervene on M (see Zhong 2020: 304). On this approach, Zhong argues, M might cause P^* in some cases, and P might cause P^* in some cases, but there is no scenario where both M and P cause P^* (311–313). Unfortunately, as Kroedel argues, “Zhong’s suggestion ultimately requires ad hoc exemptions from the holding-fixed requirement” (2479).⁴ I cannot here enter into, much less legislate, this internecine dispute. Luckily, I don’t need to. Given all these existing and as-yet-unresolved problems with variations on the theme of a counterfactual account of causation, that a specific, non-orthodox implementation of such an account is compatible with *Closure** counts for little (if anything) on the scale of plausibility in *Closure**’s favor.

Third, in addition to the previous pile of complaints, I here want to raise a new difficulty for a counterfactual account of the sort Zhong endorses: namely, that such an account undercuts there being a systematic connection between scientific laws (or “laws of nature”) and causes, of the sort which reflects a common understanding of such laws, including physical laws, as typically expressing how systems evolve over time in response to causal influences. Hence, for example, Kim takes for granted that “explanation and causation in a given domain depend on the presence/availability of laws” (2005, 337), and suggests that it is part of the physicalist world-view that “All causal processes, and explanations that track them, are fundamentally physical, in the sense that they are governed and explained by strict physical laws. In this sense, all causation and explanation are at bottom physical causation and explanation” (337-8).⁵

⁴ More specifically, Kroedel argues that the requirement of Zhong’s approach that individual supervenience bases are represented by distinct variables introduces new necessary connections (e.g., exclusionary necessary connections between incompatible realizers or determinates) of the sort making interventions impossible, where “No principled way of allowing only some variables to vary that stand in necessary connections to the variable that is intervened on is forthcoming” (2496).

⁵ The common supposition is operative across a wide range of views about causation, laws, and scientific explanation. Hence, e.g., Handfield (2009) observes that Lewis (a neo-Humean about causes and laws) and

Note that the common understanding of scientific laws that I'm driving at here does not require that scientific laws always express causal relations; only that they sometimes do so; moreover, for present purposes we may take the supposition to be neutral on whether laws are prior to causes or *vice versa* (or neither). To be sure, some reject that physical laws ever express causal relations—notably, Russell (1912) and Field (2003)—on grounds, first, that the laws of physics do not explicitly refer to causes, and second, that it does not make sense to see them as doing so.⁶ I have previously argued that the considerations raised by Russell and Field are unconvincing.⁷ In any case, those (like Zhong) who endorse a non-contrastive counterfactual account of causation are presumably committed to there being physical causation, if only in cases of “realizer indifference” (per Zhong's discussion, above). But now observe that on such a counterfactual account, whether a given physical law connecting goings-on A and B expresses that (if instantiated) A (or a token thereof) is a cause of B (or a token thereof) will be contingent on external facts or features of the situation, outside the scope of the (physical) law, including facts about whether A is a realizer of some multiply realizable macro-phenomenon C which is correlated with B in a ‘realization-insensitive way’. In other words, given a physical law connecting goings-on A and B, one will not be in a position to know whether or not A is an (actual or potential) cause of B. Hence, and more generally, such a counterfactual account of causation prevents one from systematically reading off causes from associated scientific (in particular, physical) laws, which strikes me as a bad result.⁸

Relatedly, it is unclear how Zhong's preferred account of causation can make sense of the original empirical motivations for *Closure*—namely, that seeming physically acceptable (mechanical, muscular) causes (forces/interactions) are ultimately based in physical causes (forces/interactions). My own view is that there is no problem with both realized and realizing goings-on (whether forces/interactions or associated entities, features, states of affairs) being causes of a single physical effect. But independent of whether one aims to preserve higher-level causal efficacy or relevance, it is ultimately the purported empirical fact that fundamental physical forces/interactions and laws suffice to explain and predict the evolution of physical systems, along with the common

Armstrong (a non-Humean about these notions) agree that “wherever laws are at work—at least in typical cases—there will be causation” (Handfield 2009, 15); similarly for, e.g., dispositional essentialist accounts on which scientific laws typically express causal dispositions. (The “typical” qualifications reflect that some scientific laws express constraints—symmetry principles, conservation laws—rather than causal relations or dispositions/powers per se.) The common supposition is also operative in many accounts of scientific explanation as ultimately causal explanation (for discussion see Ross and Woodward 2023 and Skow 2014). And so on.

⁶ As Field describes Russell's (and his) general line of thought: “not only does the word ‘cause’ not appear in the advanced sciences, but the laws that these sciences state are incompatible with causation as we normally understand it” (435).

⁷ In brief: that physical equations are expressed in broadly mathematical rather than explicitly causal terms plausibly reflects the usual scientific focus on the prediction of quantities, coupled with a kind of expressive underdetermination associated with logical and mathematical expressions, as opposed to encoding that physical goings-on do not enter into causal relations. Moreover, Russell's and Field's arguments for the laws of physics being “incompatible with causation as we normally understand it” ultimately rely on a certain implausible conception of causation, according to which true causal claims take the form of universal generalizations. See Wilson (2021), 34–5 for further discussion.

⁸ I leave aside here whether and if so how a contrastive counterfactual account might address this concern.

supposition that scientific laws systematically express causal relations, that provides the basis for taking any physical effect to have a physical—and not just a physically acceptable—cause. From this perspective, a non-contrastive counterfactual account of causation is implausible, and considerably so, in failing to accommodate this empirical data.

Let's sum up the results of this section. Zhong claimed that *Closure** in the new causal argument is more plausible than *Closure* in the original causal argument, on grounds that, first, *Closure* is motivated by a problematic physical (transference, conserved quantity) account of causation, and second, that *Closure*, unlike *Closure**, is likely false by lights of a counterfactual account of causation. But *Closure* is not motivated by a physical quantity account, but rather, most crucially, by inductive empirical results indicating that higher-level (including non-physical but physically acceptable) forces/interactions are reducible to fundamental physical forces/interactions. Nor does attention to a counterfactual account of causation undercut the support for *Closure*—first, because a contrastive counterfactual account is compatible with *Closure*; second, because the plethora of difficulties incurred by counterfactual accounts—including the new difficulty I have raised for Zhong's preferred counterfactual account, according to which such an account undercuts the systematic connection between causes and laws—push against such an approach to causation; and third (and relatedly), because Zhong's preferred account cannot accommodate the actual empirical support for *Closure*.

3 Can a physicalist replace *Closure* with *Closure**?

Zhong goes on to consider whether the rejection of *Closure* is incompatible with physicalism, as has been frequently maintained. As Gibb (2012) puts it:

[The physicalist] cannot deny *Closure*. To do so is to maintain that not every physical effect can be accounted for in physical terms, that something non-physical is needed to provide the complete causal story. This is not something that any physicalist would want to say. (33)

(See also Wilson 2021, 81, and many others). According to Zhong, *Closure* is undermotivated as a concomitant of physicalism, and he speculatively suggests that the supposition to the contrary reflects a conflation of ontological (metaphysical) and causal dependence:

Although philosophers say surprisingly little about why physicalism is committed to *Closure*, the basic idea seems to be clear. Physicalism is an expression of “physical hegemony”: the physical domain depends on nothing but grounds everything else. But if *Closure* is false, that means that physical entities are (sometimes) dependent on non-physical entities. This would violate the supremacy of the physical domain.

However, on a closer look, this is deeply puzzling. Physicalism is supposed to be a thesis about ontological dependence rather than causal dependence. While causal dependence is a “horizontal” relationship, ontological dependence is a

“vertical” relation in the sense that the grounding is more fundamental than the grounded. My existence causally depends on the existence of my parents, but I don’t ontologically depend on them—my parents are no more ontologically fundamental than I am. Physicalism is often metaphorically characterized as the view that all other entities “arise” out of the physical domain. This is a picture of ontological/vertical dependence. So, why does it matter to the truth of physicalism whether some physical effects have non-physical causes, if the non-physical causes are physically grounded? (7)

And again:

*Closure** can accommodate the case of downward causation in which a mental state is a non-redundant cause of the physical effect (as long as the mental cause is physically grounded), whereas such a case implies the falsity of *Closure*. (7–8)

This is an interesting line of thought, but in my view it suffers from two difficulties: first, it fails to accommodate the empirical motivations for *Closure*; second, it fails to appreciate the tensions for physicalism should *Closure* be rejected.

First, the motivation for taking *Closure* to be core to physicalism isn’t plausibly taken to reflect a conflation of horizontal and vertical notions of dependence. To be sure (as a referee pointed out), the general physicalist supposition that the physical goings-on serve as a fundamental basis for all broadly scientific phenomena is in-principle compatible with physical goings-on (sometimes or even always) not themselves being causes (hence the denial of *Closure*, so long as the physical goings-on metaphysically determine, somehow or other, whatever causation there is (including the causation of physical effects)). But even if such a conception of causation—as non-fundamental, but somehow grounded in (non-causal) physical goings-on—is compatible with physicalism per se, it is in tension with the afore-mentioned empirical support for *Closure*, whereby mechanical and other macro-forces/interactions are ontologically/metaphysically dependent on fundamental physical forces/interactions. At a minimum, Zhong owes us a story about how this empirical support might be reconceived along lines compatible with *Closure** and not *Closure*.

Second, and as the empirical support for *Closure* illustrates, metaphysical and causal dependence are not so easily prised apart. Here I want to describe another way in which metaphysical and causal dependence are deeply linked, which operates to undercut Zhong’s claim that a physicalist can accept *Closure** and the concomitant commitment to non-redundant mental causation of physical effects, “as long as the mental cause is physically grounded.”

Let us start by asking: what conditions need to be in place for a mental feature *M* of the sort at issue in Zhong’s discussion (where *M* is co-occurrent with but distinct from some physical feature *P*, and where *M* is capable of causing some physical effect *P**) to be “physically grounded”? What needs to be the case for *M* to be completely metaphysically dependent on—completely realized in, completely constituted by—physical base feature *P*?

This is the question that non-reductive physicalists are pressed to answer, and it is worth recalling certain strategies which fail to work, and why. To start, while conceptual entailment of *M* (or associated propositions) by *P* (or associated propositions)

presumably would work to establish M 's physical acceptability, such entailments are typically not in place (see, e.g., Melnyk 2008). Hence, one needs to look for some or other relation(s) as doing the trick. A modal strategy involves an appeal to supervenience, encoding abstract modal correlations of some or other strength. Here the initially promising idea was that insofar as identities (which would guarantee M 's physical acceptability, albeit in a reductive fashion) hold with metaphysical necessity, perhaps the asymmetric metaphysical supervenience of mental on physical goings-on would guarantee M 's physical acceptability in a non-reductive fashion. As it turns out, however, supervening even with metaphysical necessity is compatible with the supervenient goings-on's being clearly physically unacceptable (see, e.g., Horgan 1993, Wilson 2005, Melnyk 1999, Rosen 2010).⁹ and hopes that supervenience might provide a metaphysically neutral way of characterizing non-reductive realization remain unfulfilled. Nor will it help to add in an explanatory condition, as on Horgan's "superdupervenience," according to which the requisite supervenience relations must be "robustly ontologically explainable" (Horgan 1993). For just as causal explanation involves appeal to causal relations, "ontological explanation" presumably involves appeal to ontological relations, in which case implementing the suggestion requires specifying the ontological or metaphysical relations whose holding would explain the existence and features of M in a way making clear how and why M is physically acceptable.

The previous failures of representational, modal, and explanatory routes to ensuring the physical acceptability of features such as M led to a push, starting in the late 1970s, towards identifying substantive, broadly synchronic ("cotemporal") metaphysical relations capable of doing this work. Among the several still-live suggestions here are that features like M are functionally realized by their physical dependence base features; that features like M are determinables of physical determinates; that features like M are mereological parts, at either the type or token levels, of physical wholes; that features like M have a proper subset, at either type or token levels, of the powers of their physical dependence base features; and so on. Though controversy remains about whether one or more of these relations is operative in cases of cotemporally dependent mental features, the tolerably wide consensus appears to be that on the supposition that one or other of these relations holds between such an M and P , this holding indeed suffices to ensure that M is physically realized or (as Zhong puts it) "physically grounded."

Why is it that these seemingly diverse metaphysical relations are up to the task of establishing that a given mental feature is physically realized? To begin, it is worth registering the main threat to the physical acceptability of a mental feature such as M —namely, that M 's efficacy *vis-à-vis* P^* involves some physically unacceptable force/interaction/power/law, of the sort advanced by one or other form of strong dual-

⁹ For example, consider a version of Malbranchean occasionalism, where God brings about the same mental state upon the occurrence of the same physical state, and God is moreover consistent, doing the same in every possible world; or consider a view of properties as essentially individuated by the actual laws that govern them, such that any Strongly emergent features there might be (about which more anon) would be metaphysically necessitated by physical base features (Wilson 2005).

ism, including Cartesian dualism, British emergentism, and Moorean non-naturalism. For present purposes, it suffices to focus on the need for any formulation of physicalism to appropriately contrast with British emergentism, physicalism's best naturalistically acceptable rival. As I have previously argued (see, e.g., Wilson 1999, 2015, 2021) the (also diverse) forms of British emergentism each implement one or other form of the following schema for Strong emergence, according to which a seemingly higher-level feature is Strongly emergent just in case on at least some occasion, the feature has a new power—a power that its physical base feature does not have¹⁰:

Strong Emergence: What it is for token feature *S* to be Strongly metaphysically emergent from token feature *P* on a given occasion is for it to be the case, on that occasion, (i) that *S* cotermporally materially depends on *P*, and (ii) that *S* has at least one token power not identical with any token power of *P*.¹¹

Next, as I have again previously argued, core and common to the diverse metaphysical relations mentioned above (functional realization, the determinable/determinate relation, mereological realization, and so on) is that each guarantees satisfaction of the following condition, on any occasion of its instantiation:

Proper Subset of Powers Condition: The token powers of token feature *S*, on a given occasion, are a non-empty proper subset of the token powers of the token feature *P* on which *S* cotermporally depends, on that occasion.

More generally, approaches to non-reductive physicalism adverting to one or other specific metaphysical relation can be seen as implementing the following schema for Weak emergence¹²:

Weak Emergence: What it is for token feature *S* to be Weakly metaphysically emergent from token feature *P* on a given occasion is for it to be the case, on that occasion, (i) that *S* cotermporally materially depends on *P*, and (ii) that *S* has a non-empty proper subset of the token powers had by *P*.

¹⁰ See Wilson (2021), Ch. 2.2 for details.

¹¹ Some clarifications, which also apply to the schema for Weak emergence to be shortly discussed. First, the “cotermporal material dependence” condition encodes (i) (physical) substance monism and (ii) the assumption that emergent feature types supervene on base feature types with at least nomological necessity. Second, the notion of “power” at issue in the schema is metaphysically highly neutral: talk of powers is simply shorthand for talk of what causal contributions possession of a given feature makes (or can make, relative to the same laws of nature) to an entity's bringing about an effect, when in certain circumstances. That features are associated with actual or potential causal contributions (“powers”) reflects the uncontroversial fact that what entities do (can do, relative to the same laws of nature) depends on how they are (what features they have). Third, the base feature at issue in the schema is a feature of a given physical configuration (plurality or structural aggregate), and the novelty of the power at issue in condition (ii) should be understood accordingly. In particular, the novelty of this power is not to be understood as relative to the powers of features of individual entities entering into the dependence base configuration. Fourth, *S*'s novel power is not just nonfundamentally novel (as reflecting a merely aggregative difference between powers of a configuration and powers of components of the configuration) but is moreover fundamentally novel. I do not include the qualifier “fundamentally” in the schema because including it would misleadingly suggest that the nonfundamentality of *S*'s novel power was among the available options, which it isn't, under the operative understanding of *P* as having any nonfundamental powers attaching to a mere aggregation of lower-level components.

¹² See Wilson (2021), Ch. 2.3 for details.

Satisfaction of the *Proper Subset of Powers Condition* by a feature such as *M* clearly blocks compatibility with Strong emergentism.¹³ Moreover, *M*'s satisfaction of the condition is compatible with *M*'s distinctness from *P* (since *M* and *P* have different sets of powers); with *M*'s efficacy *vis-à-vis* *P** not involving overdetermination of a problematic “double-rock-throw” variety (since on any given occasion of causation, only one token power is exercised, not two); and with *M*'s nonetheless being distinctively efficacious (in that *M*'s distinctive set of powers encodes difference-making conditions—if my thirst *M* had been differently realized, I would have still reached for the Fresca—and/or levels of causal grain abstracting from the micro-physical details). Satisfaction of the *Proper Subset of Powers Condition* also provides a metaphysical basis for making sense of the common supposition that mental features are “multiply realized,” given that the powers of a feature such as *M* are properly seen as subsets of the powers of multiple physical base features.

Summing up so far: the diverse accounts of non-reductive physicalism appealing to specific metaphysical relations all provide an in-principle illuminating basis for establishing that a feature *M* of the sort at issue in Zhong's discussion is physically acceptable—that is, is “physically grounded.” And they do this by positing relations which plausibly guarantee that every token power of such a feature *M* is identical with a token power of its physical realizer *P*, on any given occasion.

Are there exceptions to this rule? One might wonder whether the comparatively recently introduced “Grounding,” understood as a kind of primitive generic relation or notion operative in any context where metaphysical dependence is at issue,¹⁴ might provide a basis for ensuring the physical acceptability of broadly scientific features like *M* in a way not requiring satisfaction of the *Proper Subset of Powers Condition*. But as I have argued (Wilson 2014b and 2018), Grounding qua abstract generic primitive cannot do this work, since a given posit of Grounding fails to answer any questions—even about whether Grounded goings-on exist, much less about whether they are causally efficacious, and if so how their efficacy stands to the efficacy of what Grounds them. Relatedly, unlike claims about the specific non-reductive realization relations, claims about the holding of Grounding relations (and notwithstanding common characterization of Grounding as a relation of “metaphysical explanation”) do not in fact offer any explanatory traction about the goings-on related by Grounding, and indeed introduce new questions of the “what Grounds Grounding” variety (see, e.g., Wilson 2018, Melnyk 2016). The bottom line here is that there is no generic primitive shortcut to physical acceptability. Any metaphysical illumination must come from the specific metaphysical relations—what I have called “small-g” grounding relations—of the sort that are up to the task of ensuring that features like *M* are distinct from their base features *P*, are unproblematically and distinctively efficacious *vis-à-vis* some physical *P**, are multiply realizable, and so on. And as above, that these small-g relations can

¹³ As I discuss in my 1999, satisfaction of the *Proper Subset of Powers Condition* also blocks compatibility with other of physicalism's strong dualist rivals. For example, given that the base feature is physical, satisfaction of this condition ensures that none of the powers of the cotemporally dependent feature are non-natural.

¹⁴ “Grounding should [...] be taken as primitive, as per the neo-Aristotelian approach (cf. Fine 2001, 1). Grounding is an unanalyzable but needed notion—it is the primitive structuring conception of metaphysics” (Schaffer 2009, 364).

do all this good work is ultimately a function of their satisfying the proper subset condition on token causal powers at the heart of the schema for Weak emergence.

Now return to Zhong's claim that a physicalist can accept *Closure**, and the associated possibility of non-physical non-redundant causes of physical effects, so long as these causes are "physically grounded." Recall that to say that a feature has power is just to say that the feature can, in appropriate circumstances, enter into causing a given effect. Now, on Zhong's view, one should endorse *Closure** rather than *Closure* precisely to make room for *M*, but not *P*, to cause physical effect *P**. So, on his view, *M* has a power that *P* does not have. It follows that Zhong cannot avail himself of the only working strategy for establishing that *M* is physically grounded—that is, by positing some or other relation ensuring that every power of *M* is token-identical to a power of *P*. So how does he propose that the physicalist who accepts *Closure** rather than *Closure* establish that *M* is physically acceptable? The question is pressing, not just because on Zhong's proposal, the usual routes to *M*'s physical acceptability are blocked, but because the having of a new power is the mark of Strong emergence, not Weak emergence of the non-reductive physicalist variety.

One response Zhong might offer would be to say that even if *P* doesn't have the power to cause *P**, *M*'s power to do so is not "fundamentally" novel, but is somehow constituted by *P*'s powers. Such a line might be motivated by observing that the core concern of the physicalist is to deny that there are any *fundamentally* novel forces/interactions/powers/laws. If *M*'s powers are only non-fundamentally novel vis-à-vis *P*'s, then perhaps this constitutive or causal relation between powers would provide a basis for *M*'s being physically acceptable. Alternatively, Zhong might maintain (as per his 2019, 43) that the causal powers of *M*, while not constituted or grounded by the causal powers of their physical realizers, are nonetheless constituted or grounded by (non-causal) physical facts. So far, though, it is unclear how these promissory notes are to be filled in—but filling these in is crucial, if Zhong's new causal argument is to be, as advertised, an argument compatible with a distinctively non-reductive version of physicalism. (And as above, filling in is also required to explain—more specifically: explain away—the seeming empirical support for *Closure*. No doubt Zhong has something characteristically nuanced and creative to say about these issues. Even so, it remains that any such routes to the physical acceptability of non-physical causes of physical effects will depart from the standard non-reductivist route, which proceeds, compatible with *Closure*, by way of the *Proper Subset of Powers Condition*. Hence, to the extent that Zhong's causal argument offers a new route to non-reductive physicalism, this will moreover be for a new version of this view.

Acknowledgements Thanks to Lei Zhong, Thomas Kroedel, and an anonymous referee for this journal for helpful comments on previous versions of this paper.

Data availability The author registers that this paper has no associated data.

Declarations

Ethics approval The author registers that the research did not involve humans or animals.

Competing interests The author declares no competing interests.

References

- Balog, K. (2025). The rise and fall of the mind-body problem. *Three Quarks Daily*.
- Baumgartner, M. (2009). Interventionist causal exclusion and non-reductive physicalism. *International Studies in the Philosophy of Science*, 23, 161–178.
- Bennett, J. (1987). Event causation: The counterfactual analysis. *Philosophical Perspectives*, 1, 367–386.
- Bennett, K. (2003). Why the exclusion problem seems intractable and how, just maybe, to tract it. *Noûs*, 37, 471–97.
- Clapp, L. (2001). Disjunctive properties: Multiple realizations. *Journal of Philosophy*, 98, 111–136.
- Dowe, P. (2000). *Physical causation*. Cambridge: Cambridge University Press.
- Elga, A. (2000). Statistical mechanics and the asymmetry of counterfactual dependence. *Philosophy of Science*, 68, 313–324.
- Field, H. (2003). Causation in a physical world. In M. Loux & D. Zimmerman (Eds.), *The Oxford handbook of metaphysics*. Oxford: Oxford University Press.
- Fine, K. (1975). Critical notice of Lewis, *Counterfactuals*. *Mind*, 84, 451–458.
- Fine, K. (2001). The question of realism. Philosophers'. *Imprint*, 1, 1–30.
- Gibb, S. (2010). Closure principles and the laws of conservation of energy and momentum. *Dialectica*, 64, 363–384.
- Gibb, S. (2012). Nonreductive physicalism and the problem of strong closure. *American Philosophical Quarterly*, 49, 29–42.
- Hall, N. (2004). Two concepts of causation. In J. Collins, N. Hall, & L. Paul (Eds.), *Causation and counterfactuals* (pp. 225–276). The MIT Press.
- Handfield, T, ed. (2009). *Dispositions and causes*. New York : Oxford University Press.; Clarendon Press.
- Hiddleston, E. (2005). A causal theory of counterfactuals. *Noûs*, 39, 632–657.
- Hitchcock, C. R. (1996). The role of contrast in causal and explanatory claims. *Synthese*, 107, 395–419.
- Horgan, T., (1989). Mental quausation. *Philosophical Perspectives 3: Philosophy of Mind and Action Theory*, 47–76.
- Horgan, T. (1993). From supervenience to superdupervenience: Meeting the demands of a material world. *Mind*, 102, 555–586.
- Kim, J. (1973). Causes and counterfactuals. *The Journal of Philosophy*, 70, 570–2.
- Kim, J. (1989). The myth of nonreductive materialism. *Proceedings and Addresses of the American Philosophical Association*, 63, 31–47.
- Kim, J. (1998). *Mind in a physical world*. Cambridge: MIT Press.
- Kim, J. (2005). Laws, causation, and explanation in the special sciences. *History and Philosophy in the Life Sciences*, 27, 325–338.
- Kroedel, T. (2024). How not to intervene on mental causes. *Philosophical Studies*, 181.
- LePore, E., & Loewer, B. (1987). Mind matters. *The Journal of Philosophy*, 84, 630–642.
- LePore, E., & Loewer, B. M. (1989). More on making mind matter. *Philosophical Topics*, 17, 175–91.
- List, C., & Menzies, P. (2009). Nonreductive physicalism and the limits of the exclusion principle. *Journal of Philosophy*, 106, 475–502.
- Melnyk, A. (1999). Supercalifragilisticexpialidocious. *Noûs*, 33, 144–54.
- Melnyk, A. (2008). Conceptual and linguistic analysis: A two-step program. *Noûs*, 42, 267–291.
- Melnyk, A. (2016). Grounding and the formulation of physicalism. In *Scientific Composition and Meta-physical Grounding*, (vol. 131). London: Palgrave-Macmillan.
- Menzies, P. (1989). Probabilistic causation and causal processes: A critique of Lewis. *Philosophy of Science*, 56, 642–663.
- Menzies, P., & Beebee, H. (2024). Counterfactual theories of causation. In E. N. Zalta, & Nodelman, U., (eds.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2024 edition.
- Papineau, D. (2001). The rise of physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and its discontents* (pp. 3–36). Cambridge: Cambridge University Press.
- Rosen, G. (2010). Metaphysical dependence: Grounding and reduction. In B. Hale & A. Hoffmann (Eds.), *Modality: Metaphysics, logic, and epistemology* (pp. 109–36). Oxford: Oxford University Press.

- Ross, L., & Woodward, J. (2023). Causal approaches to scientific explanation. In E. N. Zalta, & Nodelman, U. (eds.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2023 edition.
- Russell, B. (1912). On the notion of cause. *Proceedings of the Aristotelian Society*, 13, 1–26.
- Salmon, W. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.
- Schaffer, J. (2009). On what grounds what. In D. Manley, D. Chalmers, & R. Wasserman (Eds.), *Metametaphysics: New essays on the foundations of ontology* (pp. 347–383). OUP.
- Shoemaker, S. (2001). Realization and mental causation. In *Proceedings of the 20th World Congress in Philosophy*, (pp. 23–33). Cambridge: Philosophy Documentation Center. Published in revised form in Gillett and Loewer 2001, (pp. 74–98).
- Skow, B. (2014). Are there non-causal explanations (of particular events)? *British Journal for the Philosophy of Science*, 65, 445–467.
- Wilson, J. M. (1999). How superduper does a physicalist supervenience need to be? *The Philosophical Quarterly*, 49, 33–52.
- Wilson, J. M. (2005). Supervenience-based formulations of physicalism. *Noûs*, 39, 426–59.
- Wilson, J. M. (2014a). Hume's dictum and the asymmetry of counterfactual dependence. In A. Wilson, (ed.), *Chance and Temporal Asymmetry*, (pp. 258–279).
- Wilson, J. M. (2014). No work for a theory of grounding. *Inquiry*, 57, 1–45.
- Wilson, J. M. (2015). Metaphysical emergence: Weak and strong. In T. Bigaj, & C. Wüthrich, (eds.), *Metaphysical Emergence in Contemporary Physics; Poznan Studies in the Philosophy of the Sciences and the Humanities*, (pp. 251–306).
- Wilson, J. M. (2018). Grounding-based formulations of physicalism. *Topoi*, 37, 495–512. First published online in 2016)
- Wilson, J. M. (2021). *Metaphysical emergence*. Oxford: Oxford University Press.
- Woodward, J. (2015). Interventionism and causal exclusion. *Philosophy and Phenomenological Research*, 91, 303–347.
- Woodward, J. F. (2003). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.
- Zhong, L. (2019). Taking emergentism seriously. *Australasian Journal of Philosophy*, 98, 31–46.
- Zhong, L. (2020). Intervention, fixation, and supervenient causation. *Journal of Philosophy*, 117, 293–314.
- Zhong, L. (2023). A causal argument for physicalism. *Asian Journal of Philosophy*, 2, 1–9.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.