

INTERVENTION, FIXATION, AND SUPERVENIENT CAUSATION*

(Forthcoming in *The Journal of Philosophy*)

Lei Zhong

The Chinese University of Hong Kong

Interventionist theories of causation have gained wide recognition in the past decade. Many philosophers are attracted to the interventionist slogan: “No causation without manipulability, no manipulability without causation”.¹ Roughly speaking, on an interventionist account, X is a (type-level) cause of Y with respect to a variable set V if and only if an intervention that changes the value of X would also change the value of Y when all other relevant variables in V are held fixed at some value.² The interventionist approach captures an important difference between genuine causation and mere correlation: if X causes Y, a proper intervention that changes X would also change Y; if X is merely correlated with Y, Y would not change under suitable manipulation of X.³

* This work was supported by the General Research Fund of Hong Kong Research Grants Council (No. 14613516). I am grateful to an anonymous referee for this JOURNAL for helpful comments on an earlier version of this article. Thanks also to Xin Dong, Chang Liu, Yuan Ren, Jiji Zhang, and the audiences of my talks at Renmin University and Sun Yat-Sen University for fruitful discussion.

¹ Carl Gillett, “Scientific Emergentism and Its Move beyond (Direct) Downward Causation,” in Michele Paoletti and Francesco Orilia, eds., *Philosophical and Scientific Perspectives on Downward Causation* (New York: Routledge, 2017), pp. 242–62, at p. 255. See also James Woodward, *Making Things Happen: A Theory of Causal Explanation* (New York: Oxford University Press, 2003), p. 61.

² See Christopher Hitchcock, “The Intransitivity of Causation Revealed in Equations and Graphs,” this JOURNAL, XCVIII, 6 (June 2001): 273–99; Christopher Hitchcock, “A Tale of Two Effects,” *The Philosophical Review*, CX, 3 (July 2001): 361–96; Judea Pearl, *Causality: Models, Reasoning, and Influence* (New York: Cambridge University Press, 2000); Lawrence Shapiro, “Lessons from Causal Exclusion,” *Philosophy and Phenomenological Research*, LXXXI, 3 (November 2010): 594–604; Lawrence Shapiro, “Mental Manipulations and the Problem of Causal Exclusion,” *Australasian Journal of Philosophy*, XC, 3 (September 2012): 507–24; Woodward, *Making Things Happen*, *op. cit.*; and James Woodward, “Interventionism and Causal Exclusion,” *Philosophy and Phenomenological Research*, XCI, 2 (September 2015): 303–47.

³ It is important to note that interventionism should not be regarded as a second-order theory of causation that aims to define causation in terms of intervention. Since ‘intervention’ itself is a causal notion, a second-order version of interventionism would involve circularity. Hence interventionism should be treated as a first-order theory of causation. A first-order version of interventionism is non-circular and informative in that it aims to determine whether X causes Y by making use of *other* sorts of causal knowledge. See James Woodward, “Causation and Manipulability,” in Edward Zalta, ed., *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), URL = <<https://plato.stanford.edu/archives/win2016/entries/causation-mani/>>.

Interventionism enjoys some advantages over other difference-making theories of causation, such as the Lewisian counterfactual theory.⁴ An intervention-based account of counterfactuals seems to provide a better standard for characterizing causation than does the world-similarity approach. For example, interventionism can offer a more satisfactory solution to well-known problems with traditional counterfactual theories, such as the preemption problem and the overdetermination problem.⁵ Moreover, I will argue, an interventionist account can better deal with supervenient causation than Lewis's counterfactualism.

A growing number of philosophers are bringing interventionism into the field of supervenient causation (and mental causation in particular). In the present context, we are concerned with whether supervenient/subvenient *properties* are causally efficacious, whereas interventionist theories typically regard *variables* as the relata of causal claims. It is harmless, however, to assume that properties are binary variables, with the values corresponding to the *presence* or *absence* of those properties.⁶ Thus, an interventionist account of property causation can be formulated as follows:

- (N) A property X causes another property Y on a particular occasion if and only if
- (N1) If an intervention that sets X = *present* were to occur, then Y = *present*;
 - (N2) If an intervention that sets X = *absent* were to occur, then Y = *absent*.⁷

⁴ David Lewis, "Causation," this JOURNAL, LXX, 17 (Oct. 11, 1973): 556–67; and David Lewis, "Causation as Influence," this JOURNAL, XCVII, 4 (April 2000): 182–97.

⁵ See Hitchcock, "The Intransitivity of Causation Revealed in Equations and Graphs," *op. cit.*; and Woodward, *Making Things Happen*, *op. cit.*

⁶ Hitchcock, "The Intransitivity of Causation Revealed in Equations and Graphs," *op. cit.*; Woodward, *Making Things Happen*, *op. cit.*; and James Woodward, "Mental Causation and Neural Mechanisms," in Jakob Hohwy and Jesper Kallestrup, eds., *Being Reduced: New Essays on Reduction, Explanation, and Causation* (Oxford: Oxford University Press, 2008), pp. 218–62.

⁷ See Woodward, "Mental Causation and Neural Mechanisms," *op. cit.*; Christian List and Peter Menzies, "Nonreductive Physicalism and the Limits of the Exclusion Principle," this JOURNAL, CVI, 9 (September 2009): 475–502; Shapiro, "Mental Manipulations and the Problem of Causal Exclusion," *op. cit.*; and Lei Zhong, "Sophisticated Exclusion and Sophisticated Causation," this JOURNAL, CXI, 7 (July 2014): 341–60.

Let us call (N1) the ‘presence condition’, and (N2) the ‘absence condition’.⁸ When (N1) and (N2) hold, we can say that an intervention that changes the values of X would change the values of Y accordingly. Within an interventionist framework, this is equivalent to saying that property X causes (or is causally relevant to) property Y.

In discussing supervenient causation, we can, for the sake of simplicity, consider the variable set to consist of only three properties that are actually instantiated on a particular occasion: the supervenient property M, the subvenient property P, and the effect property E.⁹ See Figure 1 below:

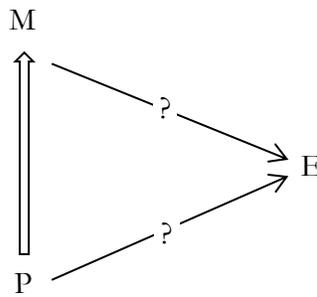


Figure 1: Supervenient Causation

Here I wish to stress that multiple realizability is commonly accepted in the literature on supervenient causation. On this view, P is only one subvenient property or realizer of M; M is also realizable by other lower-level properties, P*, P**, and so on (but they are not actually instantiated on this particular occasion).¹⁰ In the context of supervenient causation, the supervenient property and the subvenient property are assumed to be two distinct properties—otherwise it would be pointless to discuss whether the supervenient property

⁸ As I will discuss later, the two conditions should be interpreted in a proper, non-trivial way. This is relevant to how we understand intervention and fixation in the context of supervenient causation.

⁹ Many philosophers equate actual causation with token causation. But this is misleading, to say the least. Suppose that property M is causally relevant to the occurrence of E on an actual occasion *t*. This is a case of actual causation, and also a case of type causation.

¹⁰ I use the term ‘realizers’ in a minimal sense, interchangeable with ‘subvenient properties’. To say that P is a realizer (or subvenient property) of M is, in my terminology, to say that if a system *S* instantiates P, then necessarily *S* instantiates M.

can have causal efficacy apart from the causal powers of its subvenient property. And a primary reason why higher-level properties (such as mental properties and biological properties) are non-identical with their subvenient properties is that the former are multiply realizable by the latter. If M were singly realizable by and hence necessarily coextensive with P, it would be difficult to maintain that M and P are two distinct properties.

Within the interventionist framework, to decide whether M (or P) is the cause of E, we should see whether the values of E would change when an intervention changes the values of M (or P).¹¹ But the crucial question is: What kind of intervention is suitable in the case of supervenient causation? Specifically, should we hold P (or M) fixed while intervening on M (or P), and, if yes, in what way?

In the literature, many argue that interventionist supervenient causation is exempted from the fixability condition. However, this approach looks *ad hoc*, inconsistent with the general interventionist requirement on fixation. Moreover, it leads to false judgments about the causal efficacy of supervenient/subvenient properties. This article aims to clear up some common misunderstandings and to develop a novel interventionist account of supervenient causation. I argue that a suitable intervention on supervenient/subvenient properties should hold the other variables fixed at some value. The treatment of intervention and fixation that I propose can accommodate the proportionality constraint on causation and deliver correct causal verdicts in some classic examples. It is also worth noting that this interventionist account offers a promising defense of mental causation without postulating mental-physical overdetermination.

¹¹ To be more accurate, we are deciding whether M (or P) is the *direct* cause of E. Woodward makes a distinction between direct causes and contributing causes, and defines the latter in terms of the former. See Woodward, *Making Things Happen*, *op. cit.*, p. 59. But this distinction does not matter in the current context.

I. SUPERVENIENT CAUSATION AND THE FIXABILITY CONDITION

According to the classic version of interventionism, in deciding whether X is a (direct) cause of Y with respect to a variable set V , we should consider an intervention that changes X and at the same time holds fixed *all other* variables in V .¹² But classic interventionism did not intend to accommodate cases of supervenient causation. Given the supervenience relationship, it seems impossible to perform a suitable intervention on a supervenient property that could at the same time hold fixed its subvenient base. If so, it would follow that a supervenient property can never cause anything.¹³ This is what I call the *Fixability Argument* against supervenient causation:

1. There is no suitable intervention on M while holding P fixed;
2. If there is no suitable intervention on M while holding P fixed, M is not a cause of E;
3. Therefore, M is not a cause of E.

We do not want to accept the conclusion. A theory of causation should allow the possibility that supervenient properties can exert causal powers on some occasions. If any account of causation *conceptually* rules out the causal efficacy of supervenient properties, that is simply a *reductio* of the theory. To save supervenient causation, one needs to deny either Premise 1 or 2. It is striking to see that most interventionists take Premise 1 for granted and attempt to reject Premise 2. They thus later revise or develop their interventionist accounts to accommodate supervenient causation—on those updated versions, it is no longer required that P be held fixed while intervening on M. James Woodward puts it this way:

[W]hen an intervention occurs on X, its supervenience base SB(X) should not be regarded as one of those “off route causes” in IV that one needs to control for or hold fixed in intervening on X... In assessing whether X is a direct cause

¹² *Ibid.*

¹³ Michael Baumgartner, “Interventionism and Epiphenomenalism,” *Canadian Journal of Philosophy*, xI, 3 (September 2010): 359–83.

of Y, the “other variables” in V that we should hold fixed independently of the intervention on X should not include the supervenience base for X.¹⁴

And Lawrence Shapiro holds a similar view:

When investigating whether a supervening property is a cause, one must not ask whether the supervening property has causal influence *in addition* to the causal influence of its base. This question suggests the wrong kind of test, i.e. a test in which the base is held fixed while the supervening property is changed.¹⁵

However, this popular approach incurs theoretical costs. The exemption of fixability seems *ad hoc*, and is inconsistent with the general interventionist framework. Moreover, as I will argue later, this approach leads to false judgments on the causal efficacy of supervenient/subvenient properties. In contrast, I adopt an unpopular but more promising approach, which grants Premise 2 but challenges Premise 1. I will argue, for multiple theoretical reasons, that P should be held fixed while intervening on M.

Why do those philosophers believe that there is no suitable intervention on M while holding P fixed? They may have two reasons. First, some seem to understand P as the *whole* supervenience base of M, which consists of all subvenient properties or realizers of M. P is thus taken as a many-valued variable, and each value corresponds to an individual realizer of M. This explains why interventionists often make the following claims: “every change in the values of M is *necessarily* accompanied by a change in the values of P”;¹⁶ “because M *supervenes* on P, changing M while holding P fixed is *impossible*”;¹⁷ “for all values of M, the value of M *cannot* change without a change in the value of P”;¹⁸ and so on. Certainly, if P were the whole supervenience base of M, it would be *metaphysically impossible* to hold P fixed while intervening

¹⁴ Woodward, “Interventionism and Causal Exclusion,” *op. cit.*, pp. 333–4.

¹⁵ Shapiro, “Lessons from Causal Exclusion,” *op. cit.*, p. 601.

¹⁶ Baumgartner, “Interventionism and Epiphenomenalism,” *op. cit.*, p. 369.

¹⁷ Shapiro, “Lessons from Causal Exclusion,” *op. cit.*, p. 600.

¹⁸ Woodward, “Interventionism and Causal Exclusion,” *op. cit.*, p. 309.

on M. On the one hand, if M were absent under any intervention whatsoever, P would also be absent. On the other hand, if M were present under an intervention, then P would be present as well.

Nevertheless, it is problematic to understand P as the whole supervenience base of M, which is in fact an indefinite or even infinite disjunction of diverse subvenient properties. Many philosophers argue that such a disjunctive “property” is not a genuine property.¹⁹ What is more important is that, regardless of whether they are real properties, such disjunctive properties are unsuitable relata for any sort of causal relation.²⁰ According to our common understanding of causation, an event *c* causes another event *e* in virtue of certain *individual* properties that are instantiated by *c* and *e*, but not in virtue of some disjunctive properties the disjuncts of which include those individual properties. Hence the problem of supervenient causation is concerned with whether a supervenient property can exert distinctive causal powers over and above their individual subvenient properties on at least some occasions.

Therefore, P should be understood as an individual subvenient property or realizer of M. If P is only an individual subvenient property of M, it is not the case that every time the values of M change, the values of P also change—strictly speaking, M does not *supervene* upon the individual realizer P (supposing that M is multiply realizable).²¹ So, it is possible to intervene on M while holding fixed P at some value, that is, *absence*. While holding fixed P = *absent*, we can set M to be either *present* or *absent* (given multiple realizability). At first sight,

¹⁹ See, for example, D. M. Armstrong, *A Theory of Universals: Universals and Scientific Realism* (Cambridge, UK: Cambridge University Press, 1978); and Paul Audi, “How to Rule out Disjunctive Properties,” *Noûs*, XLVII, 4 (December 2013): 748–66.

²⁰ See David Lewis, “New Work for a Theory of Universals,” *Australasian Journal of Philosophy*, LXI, 4 (December 1983): 343–77.

²¹ If F supervenes on G, then there cannot be an F-difference without a G-difference. But if P is only one of the multiple realizers of M, there could be an M-difference without a P-difference. For example, two systems that have no P—they are thus identical with respect to P—could be different with respect to M.

my view that it is possible to intervene on M while holding fixed its actual realizer is *compatible* with other interventionists' claim that it is impossible to intervene on M while holding fixed all the realizers. However, our disagreements lie at a deeper level. As I will discuss in the next section, different interventionist approaches on what variables should be fixed will yield different causal verdicts.

It is worth noticing that absence is the *counterfactual* value of P. We already suppose that M and P are *actually* instantiated on a given occasion. Presence is thus the actual value of P. When some philosophers talk about holding fixed other variables, they are inclined to understand this as *holding fixed other variables at their actual values*, which is a natural and default reading of fixation. Since P is a supervenient property of M, we cannot intervene on M while holding P fixed at its actual value *presence*. When an intervention sets $M = \textit{absent}$, P must take the counterfactual value *absence*. As Shapiro says, "The proposition 'P & \sim M' is impossible. P cannot be present while M is absent".²² This is the second reason why some believe that there is no suitable intervention on M while holding P fixed.

However, this reason does not withstand careful examination. Although in many cases we should hold the variables fixed at their actual values,²³ there are some cases in which the counterfactual values should be held fixed instead. Consider a case of symmetric overdetermination (see Figure 2). Suppose that two assassins' shootings result in a victim's death. The two shootings are set up in such a way that either would have killed the victim even if the other had failed. Thus, the death of the victim (D) is causally overdetermined by

²² Shapiro, "Lessons from Causal Exclusion," *op. cit.*, p. 600.

²³ Consider an example of joint causation. Suppose that the striking of a match and the presence of oxygen jointly cause the lighting of the match. While holding oxygen fixed at the counterfactual value *absence*, lighting will not occur regardless of whether striking is present or absent. This would deliver a false verdict that striking is not a cause of lighting. However, supervenient causation is clearly not a case of joint causation. For more discussion, see Jaegwon Kim, *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation* (Cambridge, MA: MIT Press, 1998); and Jaegwon Kim, *Physicalism, or Something Near Enough* (Princeton: Princeton University Press, 2005).

the two shootings (C_1 and C_2). If we hold C_1 (or C_2) fixed at the actual value *presence*, interventions on C_2 (or C_1) will not change the effect D . That would deliver a false verdict that neither C_1 nor C_2 is the cause of D . This is why some philosophers suggest that we hold C_1 (or C_2) fixed at their counterfactual values in intervening on C_2 (or C_1).²⁴ While holding C_2 fixed at the counterfactual value *absence*, $C_1 = \textit{present}$ would lead to $D = \textit{present}$, and $C_1 = \textit{absent}$ would lead to $D = \textit{absent}$. It thus follows that C_1 is a cause of D . Similar reasoning would establish that C_2 is also a cause of D .

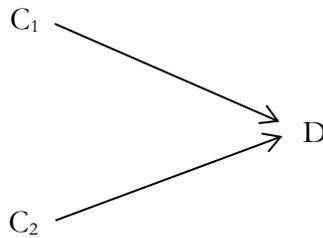


Figure 2: Symmetric Overdetermination

Since there are some cases in which we should hold the variables fixed at their counterfactual values, why not in the case of supervenient causation?²⁵ Whereas in many cases we can hold the variables fixed at either their actual values or counterfactual values, it is impossible to hold the subvenient property fixed at its actual value while intervening on the supervenient property—we can hold P fixed *only* at its counterfactual value while intervening on M . We have a *prima facie* reason to hold P fixed at the counterfactual value *absence*, because this is the only way to satisfy the general interventionist requirement that a suitable intervention should

²⁴ Hitchcock, “The Intransitivity of Causation Revealed in Equations and Graphs,” *op. cit.*; and Woodward, *Making Things Happen, op. cit.*

²⁵ For the purposes of this article, I do not have to provide a unifying account to accommodate all and only cases in which holding the variables fixed at their counterfactual values should be allowed. I only need to argue that we should hold the subvenient property fixed at its counterfactual value while intervening on the supervenient property.

hold all other variables fixed at some value. Moreover, I will argue that holding P fixed at *absence* would deliver correct and interesting verdicts on supervenient causation.

II. INTERVENTION, FIXATION, AND THE PROPORTIONALITY CONSTRAINT

Fixing P at the counterfactual value *absence* has significant theoretical advantages. This account of fixation not only respects the general requirements of interventionism, but also accommodates the *Proportionality Constraint on Causation*, according to which the cause should be proportional—neither *too general* nor *too specific*—to the effect.²⁶ Consider two examples:

[SOPHIE] The pigeon Sophie is conditioned to peck at and only at red objects. On a particular occasion, Sophie sees a scarlet (also red) object and then pecks at it. Which property, red or scarlet, is the cause of Sophie’s pecking? (Figure 3)

[ALICE] The pigeon Alice is conditioned to peck at and only at scarlet objects. On a particular occasion, Alice sees a scarlet (also red) object and then pecks at it. Which property, red or scarlet, is the cause of Alice’s pecking? (Figure 4)²⁷

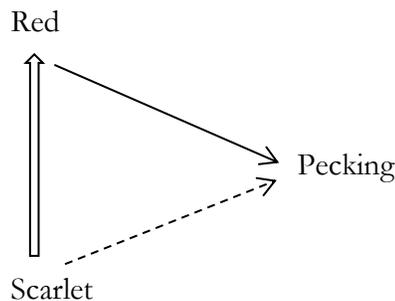


Figure 3: SOPHIE

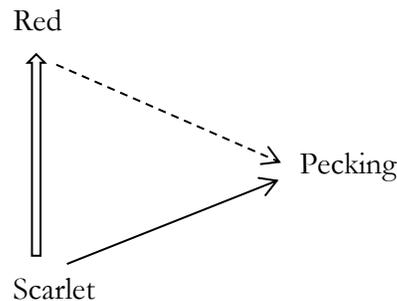


Figure 4: ALICE

On the one hand, in the Sophie case, red seems more suitable than scarlet to be the cause of pecking.²⁸ As long as an object is red, Sophie will peck at the object, regardless of whether it

²⁶ See, for example, List and Menzies, “Nonreductive Physicalism and the Limits of the Exclusion Principle,” *op. cit.*; and Stephen Yablo, “Mental Causation,” *The Philosophical Review*, CI, 2 (April 1992): 245–80.

²⁷ The two pigeon examples are borrowed respectively from Yablo, “Mental Causation,” *op. cit.*; and Sydney Shoemaker, *Identity, Cause, and Mind* (Oxford: Oxford University Press, 2003).

²⁸ Here we are concerned with which property, red or scarlet, is causally relevant to pecking. It is safe to say that red and scarlet are two different properties, regardless of whether they could be instantiated by the same event.

is scarlet. Scarlet is thus *too specific* to be the cause of Sophie’s pecking.²⁹ On the other hand, in the Alice example, scarlet rather than red should be regarded as the cause of pecking. Alice will not peck at a red object if it is not also scarlet. In other words, red is *overly general* for being the cause of Alice’s pecking.³⁰

The proportionality constraint is a reasonable idea shared by many philosophers, and the two pigeon examples give a very good illustration of this idea. The main thesis of this paper is not to argue for proportionality; that is beyond the scope of this article.³¹ In what follows, I would like to show how my particular approach to supervenient causation and the proportionality intuition can *mutually* support each other.

Let us return to the two pigeon examples. While holding fixed scarlet at the counterfactual value *absence*, we perform interventions that change the values of red. Then we see whether the values of pecking would change accordingly. Consider SOPHIE first. Within my interventionist framework, red is the cause of Sophie’s pecking if and only if the following two conditionals are true:

(S1) If an intervention sets RED = *present* (while holding SCARLET fixed at *absence*), then SOPHIE’S PECKING = *present*;

²⁹ Lewis, “Causation,” *op. cit.*; Lewis, “Causation as Influence,” *op. cit.*; List and Menzies, “Nonreductive Physicalism and the Limits of the Exclusion Principle,” *op. cit.*; Panu Raatikainen, “Causation, Exclusion, and the Special Sciences,” *Erkenntnis*, LXXIII, 3 (November 2010): 349–63; Yablo, “Mental Causation,” *op. cit.*; Zhong, “Sophisticated Exclusion and Sophisticated Causation,” *op. cit.*; and Lei Zhong, “Semantic Normativity and Semantic Causality,” *Philosophy and Phenomenological Research*, XCIV, 3 (May 2017): 626–45.

³⁰ List and Menzies, “Nonreductive Physicalism and the Limits of the Exclusion Principle,” *op. cit.*; Yablo, “Mental Causation,” *op. cit.*; Zhong, “Sophisticated Exclusion and Sophisticated Causation,” *op. cit.*; and Zhong, “Semantic Normativity and Semantic Causality,” *op. cit.*

³¹ Some philosophers maintain that proportionality only matters to explanation rather than causation. See, for example, Neil McDonnell, “Causal Exclusion and the Limits of Proportionality,” *Philosophical Studies*, CLXXIV, 6 (June 2017): 1459–74; and Brad Weslake, “Proportionality, Contrast and Explanation,” *Australasian Journal of Philosophy*, XCI, 4 (December 2013): 785–97. On their view, both red and scarlet are the cause of Sophie’s pecking, although red figures in a better causal explanation of why Sophie pecks at the object. However, the separation of causation from causal explanation is unconvincing. Facts about causal explanation are grounded in facts about causation. If red plays a genuinely better causal explanatory role than scarlet does in the SOPHIE case, this is probably because red is more suitable than scarlet to be the cause of Sophie’s pecking.

(S2) If an intervention sets RED = *absent* (while holding SCARLET fixed at *absence*), then SOPHIE'S PECKING = *absent*.³²

By assumption, Sophie is conditioned to peck at *all* red objects. That is, whenever red is present (even if the color is not scarlet), Sophie's pecking will always occur. Thus, (S1) is true. On the other hand, Sophie is conditioned to peck at *only* red objects. If an object is not red, Sophie will not peck at it. (S2) is also true. Therefore, red is the cause of Sophie's pecking.

Is red also the cause of Alice's pecking? Consider the following two conditionals:

(A1) If an intervention sets RED = *present* (while holding SCARLET fixed at *absence*), then ALICE'S PECKING = *present*;

(A2) If an intervention sets RED = *absent* (while holding SCARLET fixed at *absence*), then ALICE'S PECKING = *absent*.

Alice is conditioned to peck at only scarlet objects. If an object is not scarlet, Alice will not peck at it. So, whereas (A2) is true, (A1) is false. Suppose that Alice were to see a crimson object under interventions (this is a case in which RED = *present* but SCARLET = *absence*). In that situation, Alice would not peck at the object, even though it was red. That is to say, (A1) is not satisfied. Therefore, red does not cause Alice's pecking.

As I mentioned earlier, the interventionist account of property causation (N) consists of two conditions: the presence condition (N1) and the absence condition (N2). While some theories of causation only focus on the presence condition,³³ other theories merely

³² Someone might say that if scarlet is not on a causal path to pecking in the SOPHIE case, we do not need to hold scarlet fixed while intervening on red. However, whether the supervenient/subvenient property is causally relevant to the effect is precisely what is at issue in the supervenient causation debate. It is inappropriate to assume that scarlet does not cause Sophie's pecking, *before* we offer interventionist diagnosis. On my approach, scarlet and red are treated as competing causal candidates, and therefore one should be held fixed while intervening on the other *regardless of whether* it is SOPHIE or ALICE. But even if scarlet is held fixed while intervening on red in the SOPHIE case, a change in the values of red would still lead to a change in the values of pecking.

³³ See Phil Dowe, *Physical Causation* (New York: Cambridge University Press, 2000); and Wesley Salmon, *Scientific Explanation and the Causal Structure of the World* (Princeton: Princeton University Press, 1984).

emphasize the absence condition.³⁴ However, both conditions are significant and should be interpreted in a non-trivial way for the sake of providing a plausible account of causation. The absence condition can disallow putative causes that are not general enough, and the presence condition can rule out inadequately specific causes.³⁵ The account of intervention and fixation that I propose can provide a suitable interpretation of the two conditions.

What if we do not hold fixed scarlet while intervening on red? Given that scarlet is a subvenient property of red, the only situation in which we intervene on red *without* holding scarlet fixed is this: (1) when an intervention sets RED = *present*, the other variable SCARLET = *present*; and (2) when an intervention sets RED = *absent*, the other variable SCARLET = *absent*. If so, the two interventionist conditions would be met (regardless of whether it is the SOPHIE case or the ALICE case). First, the presence condition is satisfied in a *trivial* way. When the other variable SCARLET = *present*, RED = *present* will lead to PECKING = *present*. Second, the absence condition is also satisfied. When the other variable SCARLET = *absent*, RED = *absent* will lead to PECKING = *absent*. Thus, we would come to an absurd conclusion that red causes pecking not only in SOPHIE but also in ALICE.

So, if we do not hold fixed the subvenient property while intervening on the supervenient property, we would get an *overly optimistic* account of supervenient causation,

³⁴ See Lewis, "Causation," *op. cit.*; and Lewis, "Causation as Influence," *op. cit.* See also Thomas Kroedel, "A Simple Argument for Downward Causation," *Synthese*, CXCII, 3 (March 2015): 841–58. Lewis's world-similarity criteria are committed to *Strong Centering*, which asserts that the actual world is closer to itself than any other possible world. Given *Strong Centering*, it is trivially true that if *C* had occurred, *E* would have occurred (where *C* and *E* are two actual events or property instances). So, for Lewis, the presence condition on causation is insignificant; his theory of causation is thus primarily concerned with the absence condition "if *C* had not occurred, *E* would not have occurred." Many suggest that we should adopt a weaker centering assumption for various theoretical considerations (see, for example, List and Menzies, "Nonreductive Physicalism and the Limits of the Exclusion Principle," *op. cit.*). It is important to note that *Strong Centering* does not hold in the interventionist framework (see Woodward, "Causation and Manipulability," *op. cit.*).

³⁵ (N) does not logically entail the proportionality constraint—whether (N) delivers the proportionality constraint depends upon how we understand intervention and fixation. But interventionism seems to have more theoretical resources than other theories of causation to accommodate the proportionality constraint.

which fails to differentiate between the two kinds of cases (SOPHIE vs. ALICE). In contrast, my interventionist account of supervenient causation is neither *pessimistic* nor *cheap*.

According to my theory, supervenient properties can have causal powers in some cases but not in other cases.

III. THE CAUSAL EFFICACY OF SUBVENIENT PROPERTIES

Interventionists are mainly concerned with whether supervenient properties have causal powers. They have not explicitly discussed the causal efficacy of *subvenient* properties within the interventionist framework. Perhaps they take it for granted that subvenient properties are causally efficacious. Suppose that we are now deciding whether P is the cause of E. What kind of intervention should we adopt? Should we hold M fixed while intervening on P? Or, as in the pigeon examples, should we hold red fixed (at some value) while intervening on scarlet?

I contend that we should hold red fixed at its actual value *presence* while intervening on scarlet. Although this way of fixation—holding red fixed at the actual value—is apparently different from the proposal mentioned in the previous section (that is, holding scarlet fixed at its counterfactual value), the underlying rationale is the same. Holding red fixed at *presence* is the only way to satisfy the general interventionist requirement on fixation. If we take both the absence and presence conditions into account, red cannot be held fixed at *absence* in intervening on scarlet. Moreover, this way of intervention and fixation will deliver correct causal verdicts.

Consider ALICE first. Within my interventionist framework, scarlet is the cause of Alice's pecking if and only if the following two conditionals are true:

(A3) If an intervention sets SCARLET = *present* (while holding RED fixed at *presence*), then ALICE'S PECKING = *present*;

(A4) If an intervention sets SCARLET = *absent* (while holding RED fixed at *presence*), then ALICE'S PECKING = *absent*.

By assumption, Alice is conditioned to pick at *all* scarlet objects. That is, if scarlet is present, Alice will always peck. Thus, (A3) is true. On the other hand, Alice is conditioned to pick at *only* scarlet objects. If an object is not scarlet (even if it is still red), Alice will not peck at it. (A4) is also true. Therefore, scarlet is the cause of Alice's pecking.

Is scarlet also the cause of Sophie's pecking? Consider the following two conditionals:

(S3) If an intervention sets SCARLET = *present* (while holding RED fixed at *presence*), then SOPHIE'S PECKING = *present*;

(S4) If an intervention sets SCARLET = *absent* (while holding RED fixed at *presence*), then SOPHIE'S PECKING = *absent*.

Sophie is conditioned to peck at all red objects. As long as an object is red (regardless of whether it is scarlet), Sophie will peck at it. So, whereas (S3) is true, (S4) is false. Suppose that Sophie sees a crimson object under an intervention (this is a case in which SCARLET = *absent* but RED = *present*). In that situation, Sophie will still peck at the object, even though it is not scarlet. That is to say, the absence condition is not satisfied. Therefore, scarlet is *not* the cause of Sophie's pecking.

Some philosophers only focus on the presence condition. Whenever scarlet is present, pecking is also present, regardless of whether it is the ALICE case or the SOPHIE case. However, we should take both conditions seriously. Consider a similar example. Suppose that a palm tree would be broken if the wind speed reaches 60 MPH or more. On a particular occasion, the wind speed reaches 75 MPH, and the palm tree is broken. But which property is causally relevant: the wind's blowing at more than 60 MPH, or the wind's

blowing at 75 MPH?³⁶ It seems reasonable to say that the former rather than the latter is the cause. Although blowing at 75 MPH satisfies the presence condition, it fails to meet the absence condition. Likewise, scarlet is not the cause of Sophie's pecking, as the absence condition is not satisfied.

But, if we do *not* hold fixed red while intervening on scarlet, the absence condition would be met in a trivial way. Because scarlet is a subvenient property of red, the only way to intervene on scarlet without holding red fixed is this: (1) when an intervention sets SCARLET = *present*, the other variable RED = *present*; and (2) when an intervention sets SCARLET = *absent*, the other variable RED = *absent*. On the one hand, the presence condition would be satisfied. Supposing that the other variable RED = *present*, if SCARLET = *present*, then PECKING = *present*. On the other hand, the absence condition would be (trivially) satisfied. Given that RED = *absent*, if SCARLET = *absent*, then PECKING = *absent*. Thus, it follows that scarlet is the cause of pecking in both SOPHIE and ALICE. While intervening on red without holding fixed scarlet provides a trivial interpretation of the *presence* condition, intervening on scarlet without holding fixed red gives a trivial reading of the *absence* condition.

It is worth noticing that my interventionist explanation of why scarlet is not the cause of Sophie's pecking is different from the traditional counterfactualist approach that adopts Lewis's world-similarity criteria. Some philosophers would say that scarlet is not the cause of Sophie's pecking because the counterfactual 'if scarlet had not occurred, then Sophie would not have pecked' is false.³⁷ Why is it false? According to the counterfactualist, the closest possible world in which scarlet is absent is a world in which another shade of red (say, crimson) is somehow present, and Sophie still pecks.

³⁶ The two properties are distinct properties, although they may be instantiated by the same event (supposing that we adopt a coarse-grained conception of events).

³⁷ List and Menzies, "Nonreductive Physicalism and the Limits of the Exclusion Principle," *op. cit.*

However, it is controversial to claim that the closest possible world in which scarlet is absent must be one in which another shade of red is present. For the sake of argument, suppose that in the actual world the objects that can be presented to Sophie consist mainly of non-red (green, blue, yellow, and so forth) objects. Sophie is conditioned to pick at all and only red objects. On a particular occasion, Sophie sees a scarlet object and then pecks at it. But, if counterfactually the scarlet object did not occur, there would be a non-red object presented to Sophie. So, the closest world in which scarlet is absent is a possible world in which red is also absent. Hence it follows that the counterfactual ‘if scarlet had not occurred, then Sophie would not have pecked’ is true. But is scarlet the cause of Sophie’s pecking in this scenario? It seems not. The question of whether scarlet is the cause of Sophie’s pecking does not hinge on the truth of counterfactuals that are evaluated in terms of world similarity. Instead, we should appeal to interventionist counterfactuals like (S3) and (S4) to determine whether scarlet is the cause of pecking. If an intervention that sets scarlet to be absent (while holding red fixed) would still lead to the presence of pecking, scarlet should not be regarded as the cause of Sophie’s pecking, even though some world in which scarlet is absent and pecking is also absent is closer to the actual world than any world in which scarlet is absent and pecking is present. Thus, an interventionist account can better deal with supervenient causation than Lewis’s counterfactualism.

In the last part of this section, I want to consider a possible objection to my view. Notice that if we set scarlet to be *absent* while holding fixed red at *presence*, a new realizer of red (say, crimson) will be *present*. Some philosophers might maintain that this kind of intervention violates the original framework of interventionism. For example, Neil McDonnell writes:

According to Zhong (2014, p. 356), when an intervention sets scarlet to absent, and red is held at present, a third variable, crimson, may also change value (to

present). Crimson, in this case, is an *off-path* variable which Woodward's definition of an intervention (2003, p. 98) states should be held fixed when considering the causal role of scarlet (to rule it out as a confounding variable).³⁸

However, it is incorrect to say that crimson is a *confounding* variable. In the ALICE case, an intervention that sets scarlet = *absent* while holding red fixed at *presence* would change the values of pecking (from *presence* to *absence*), regardless of whether red is realized by crimson or ruby or vermilion... In the SOPHIE case, an intervention that sets scarlet = *absent* while holding red fixed at *presence* would not change the values of pecking, whether red is realized by crimson or not. In either case, crimson should not be considered as a confounder.

Here we are faced with two choices on intervention: (1) an intervention that makes both scarlet and red absent; or alternatively (2) an intervention that makes scarlet absent but keeps red present. Some may fail to realize that the values of some variables would change under *either* intervention. Certainly, if we perform intervention (2), then the values of crimson will change, from the actual value *absence* to the counterfactual value *presence*. But if we perform intervention (1) the values of red will also change, from the actual value *presence* to the counterfactual value *absence*. Which intervention is a suitable intervention? I think we should choose (2) for two reasons.

First, as I have argued thus far, the second kind of intervention (2) would deliver correct causal judgments. My proposed account of intervention and fixation can accommodate not only cases in which subvenient rather than supervenient properties are causally responsible for the effects (such as the ALICE case), but also cases in which supervenient rather than subvenient properties are the real causes (such as the SOPHIE case). But, if we adopt the first sort of intervention (1), it would lead to a mistaken view that subvenient properties are *always* causally efficacious.

³⁸ McDonnell, "Causal Exclusion and the Limits of Proportionality," *op. cit.*, p. 1467.

Second, intervention (2) is more coherent with the general interventionist framework than intervention (1) is. In the pigeon examples, two competing causal candidates (red and scarlet) are instantiated on a particular occasion. To determine whether one variable is the cause, we should intervene on it while holding the other variable fixed at some value. But crimson is not actually instantiated on this occasion—it is not presented in the causal graphs (Figures 3 & 4). Since we cannot hold both red and crimson fixed, holding fixed the instantiated property red should be given the priority. I admit that this is different from a case of non-supervenient causation (such as the case of symmetric overdetermination) in which intervening on one variable while holding fixed the other variable will *not* bring about a new variable. Nevertheless, given the supervenience relationship, if we set scarlet to be absent while holding red fixed at *presence*, this must make a third variable, say crimson, present. But this difference seems irrelevant to the question of which instantiated property, red or scarlet, is the cause of pecking.

IV. COMPATIBILISM, EXCLUSIONISM, AND INTERACTIONISM

A significant theoretical upshot of my interventionist account of supervenient causation is that it can shed some light on mental causation. Specifically, this account will offer a distinctive and promising solution to the *exclusion problem*, the most powerful challenge to non-reductive physicalism.³⁹ Consider a standard formulation of the Exclusion Argument:

1. [CLOSURE] Every physical effect has a sufficient physical cause (at any time at which it has a cause)—that is, every physical event is caused by another physical event solely in virtue of physical properties.
2. [EFFICACY] Some physical effects have mental causes—that is, some physical events are caused by mental events in virtue of mental properties.

³⁹ Kim, *Mind in a Physical World*, *op. cit.*; and Kim, *Physicalism, or Something Near Enough*, *op. cit.*

3. [EXCLUSION] It is not the case that mental properties and physical properties causally overdetermine the same effects in a systematic way.⁴⁰
4. [IDENTITY] Therefore, mental properties are identical with physical properties.⁴¹

Let us return to Figure 1. Suppose that M is a mental property, P is a physical subvenient property of M, and E is another physical property. According to CLOSURE, P is the cause of E. According to EFFICACY, E is also caused by M. Given EXCLUSION, M and P could not both cause E if they were distinct from each other. Thus, it follows that M is identical with P.

Most philosophers do not accept the IDENTITY conclusion and raise various objections to the exclusion argument. This argument seems valid, but is it sound? A popular solution to the exclusion problem is the so-called *Compatibilist* approach, which attempts to refute the EXCLUSION premise.⁴² In contrast, alternative approaches that accept EXCLUSION can be labelled as *Exclusionism*. Compatibilism proposes that mental and physical properties systematically overdetermine the physical effects: while P is a sufficient physical cause of E, M is another sufficient cause—or at least part of another sufficient cause—of E.⁴³

It is worth noticing that some interventionists adopt the compatibilist approach as well. For example, Shapiro says that “Because changing M is impossible without

⁴⁰ Here I use the term ‘causal overdetermination’ in a broad sense: to say that A and B causally overdetermine the effect E is *just* to say that A and B are two distinct and sufficient causes of E.

⁴¹ For the purpose of establishing a *robust* version of reductive physicalism, the exclusion argument is standardly formulated in terms of mental properties rather than mental events.

⁴² See Karen Bennett, “Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It,” *Noûs*, XXXVII, 3 (September 2003): 471–97; Barry Loewer, “Mental Causation, or Something Near Enough,” in Brian McLaughlin and Jonathan Cohen, eds., *Contemporary Debates in Philosophy of Mind* (Oxford: Blackwell, 2007), pp. 243–64; Jonathan Schaffer, “Overdetermining Causes,” *Philosophical Studies*, CXIV, 1–2 (May 2003): 23–45; Sydney Shoemaker, *Physical Realization* (Oxford: Oxford University Press, 2007); Theodore Sider, “What’s So Bad about Overdetermination?,” *Philosophy and Phenomenological Research*, LXVII, 3 (November 2003): 719–26; and Jessica Wilson, “Non-reductive Realization and the Powers-based Subset Strategy,” *The Monist*, XCIV, 1 (January 2011): 121–54.

⁴³ Mental-physical overdetermination is sometimes simplified as the case in which the mental property (M) and the physical subvenient property (P) are two distinct and sufficient causes of the same effect (E). But we should note that the compatibilist does not have to say that M *alone* is another sufficient cause of E—she only needs to make the weaker claim that M *is part of* another sufficient cause of E.

simultaneously changing M's supervenience base P, and because P is a cause of P*, a change in M *does* result in a change in P*. *This is evidence that M is a cause of P**.”⁴⁴ Woodward also contends that we should treat both the mental property M1 and the physical realizer P1 as causing another physical effect P2.⁴⁵

However, compatibilism strikes many as highly problematic. In the compatibilist picture, “mental properties look like freeloaders, merely piggybacking on the real bearers of causal powers”.⁴⁶ Given that the mental ontologically depends upon the physical, mental-physical overdetermination is a model of *dependent* overdetermination. This is different from the case of *symmetric* overdetermination in which the two sufficient causes of the effect are metaphysically *independent* of each other: even if one cause had not occurred, the other cause would have still occurred (see Figure 2).⁴⁷ Jaegwon Kim puts it this way:

In standard cases of overdetermination, like two bullets hitting the victim's heart at the same time, the short circuit and the overturned lantern causing a house fire, and so on, each overdetermining cause plays a distinct and distinctive causal role. The usual notion of overdetermination involves two or more separate and independent causal chains intersecting at a common effect. Because of supervenience, however, that is not the kind of situation we have here. In this sense, this is not a case of genuine causal overdetermination.⁴⁸

My interventionist account offers a straightforward and unifying justification of Kim's intuition that while symmetric overdetermination is intelligible, dependent overdetermination is dubious. Consider a case of symmetric overdetermination. Two shootings (C_1 and C_2) are followed by the death of the victim (D). To determine whether C_1 is a cause of D, we intervene upon C_1 while holding C_2 fixed at some value (namely, at the counterfactual value

⁴⁴ Shapiro, “Lessons from Causal Exclusion,” *op. cit.*, p. 601.

⁴⁵ Woodward, “Interventionism and Causal Exclusion,” *op. cit.*

⁴⁶ David Robb and John Heil, “Mental Causation,” in Edward Zalta, ed., *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), URL = < <https://plato.stanford.edu/archives/win2018/entries/mental-causation/> >.

⁴⁷ Some philosophers use the term ‘overdetermination’ roughly in the sense of independent overdetermination—in a narrower sense than I do. See, for example, Bennett, “Why the Exclusion Problem Seems Intractable and How, Just Maybe, to Tract It,” *op. cit.*

⁴⁸ Kim, *Physicalism, or Something Near Enough*, *op. cit.*, p. 48.

absence). Then we find that $C_1 = present$ leads to $D = present$, and $C_1 = absent$ leads to $D = absent$. The presence and absence conditions are satisfied. It thus follows that C_1 is a cause of D . A similar reasoning would establish that C_2 is also a cause of D . C_1 and C_2 causally overdetermine D , because an intervention on either C_1 or C_2 will change D (when one holds the other variables fixed).

But neither ALICE nor SOPHIE is like this. There is only one variable (red or scarlet, but not both) on which a suitable intervention will change the values of pecking. Consider ALICE first. If we intervene on scarlet while holding red fixed (at *presence*), we will find that (1) if SCARLET = *present*, then PECKING = *present*; and (2) if SCARLET = *absent*, then PECKING = *absent*. Scarlet is a cause of Alice’s pecking. On the other hand, if we intervene on red while holding scarlet fixed at *absence*, PECKING = *absent* regardless of whether RED = *present* or *absent*. Thus, red is not a cause of Alice’s pecking. While there is an intervention on scarlet that will change the values of Alice’s pecking, there is no intervention on red that can change Alice’s pecking (supposing that the other variables are held fixed). That is to say, red and scarlet do *not* causally overdetermine Alice’s pecking. Similarly, we can establish that SOPHIE is not a case of causal overdetermination either. Although a suitable intervention on red will change the values of Sophie’s pecking, no proper intervention on scarlet can change Sophie’s pecking.

Compatibilists often say that given the close relationship between the mental and the physical (supervenience, realization, determinable/determinate, and so forth), their causal powers do not compete with each other—mental properties rather “inherit” the causal efficacy of physical properties.⁴⁹ This claim is, however, unconvincing. As I have argued,

⁴⁹ See Bennett, “Why the Exclusion Problem Seems Intractable and How, Just Maybe, to Tract It,” *op. cit.*; Shoemaker, *Physical Realization*, *op. cit.*; and Wilson, “Non-reductive Realization and the Powers-based Subset Strategy,” *op. cit.*

although red and scarlet hold such a close relationship, the determinable/determinate relation, they still causally compete with each other. Either the supervenient property red is the cause (as in the SOPHIE example) or the subvenient property scarlet is the cause (as in the ALICE example), but not both—this is not a case of causal overdetermination.

Thus, an interventionist account of supervenient causation understood properly will favor exclusionism over compatibilism. But even though exclusionism is true—even though the mental and the physical causally compete with each other—there is a further question to be answered: Is the mental-physical correlation like the SOPHIE case or the ALICE case? Consider an example. I was in pain and hence started groaning. Suppose that the property of having such and such neurons firing in the anterior cortex (call this property ‘P’) is the particular realizer of my pain in this case. But there are many other similar properties that can also realize pain. Will a person still groan when her pain is realized by a different physical property? Or is the groaning behavior only sensitive to the particular realizer P?⁵⁰

Let me introduce the distinction between realization-sensitivity and realization-insensitivity.⁵¹ 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—ALICE is a realization-sensitive case. 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 small change—SOPHIE is a realization-insensitive case.

If every putative case of mental causation were a realization-sensitive case, mental properties would, on my interventionist account, have no genuine causal powers—it is rather

⁵⁰ Thanks to an anonymous reviewer for pressing me to address this issue.

⁵¹ List and Menzies, “Nonreductive Physicalism and the Limits of the Exclusion Principle,” *op. cit.*; Woodward, “Mental Causation and Neural Mechanisms,” *op. cit.*; Zhong, “Semantic Normativity and Semantic Causality,” *op. cit.*; and Lei Zhong, “Taking Emergentism Seriously,” *Australasian Journal of Philosophy*, XCVIII, 1 (March 2020): 31–46.

the physical realizers that would do the causal work. This is the picture of *Epiphenomenalism*, in which human agency is nothing but an illusion. But fortunately, the reality is far from that depressing. The question of whether mental-physical correlations are realization-sensitive or realization-insensitive is largely an empirical question. According to research in psychology and neuroscience, the connections between mental states and physical states (such as behaviors) seem to be realization-insensitive at least on some occasions.⁵² Moreover, in the literature on mental causation, no one denies that the mental-physical correlation could be realization-insensitive—we only disagree over the causal implications of realization-insensitivity. Some philosophers (mistakenly) believe that mental property M and its physical realizer P causally overdetermine the physical effect E, even when M is realization-insensitive.⁵³ In contrast, according to my interventionist approach, M rather than P should be regarded as the cause of E if it is a realization-insensitive case. Consider the following four conditionals:

- (a) If an intervention sets M = *present* (while holding P fixed at *absence*), then E = *present*;
- (b) If an intervention sets M = *absent* (while holding P fixed at *absence*), then E = *absent*.
- (c) If an intervention sets P = *present* (while holding M fixed at *presence*), then E = *present*;
- (d) If an intervention sets P = *absent* (while holding M fixed at *presence*), then E = *absent*.

Given realization-insensitivity, the conditionals (a), (b) and (c) are true, but (d) is false. Since both (a) and (b) are true, M is a cause of E. But because (d) does not hold, P is not a cause of

⁵² See, for example, Ronald Endicott, “Species-Specific Properties and More Narrow Reductive Strategies,” *Erkenntnis*, xxxviii, 3 (May 1993): 303–21; Terence Horgan, “Nonreductive Materialism and the Explanatory Autonomy of Psychology,” in Steven Wagner and Richard Warner, eds., *Naturalism: A Critical Appraisal* (Notre Dame: University of Notre Dame Press, 1993), pp. 295–320; and Bryan Kolb and Ian Whishaw, *Fundamentals of Human Neuropsychology*, 5th Edition (New York: Worth Publishers, 2003), pp. 621–41.

⁵³ Woodward, “Interventionism and Causal Exclusion,” *op. cit.*, p. 305.

E. That is, the mental property M causes the physical effect E in a non-overdetermining way—we can call it the case of *Interactionism*. Epiphenomenalism and interactionism are two types of exclusionism. But unlike epiphenomenalism, interactionism holds the view that some physical effects have mental causes (EFFICACY).

Does interactionism entail the falsity of CLOSURE? It depends. We should distinguish between two versions of CLOSURE in terms of a narrow and a broad notion of the physical. While physical properties in a narrow sense are fundamental physical properties, properties that are investigated by (ideal) physics, physical properties in a broad sense include both fundamental physical properties and higher-level physical properties (such as behavioral, physiological, and biological properties). Now we can formulate two versions of CLOSURE accordingly:

[NARROW CLOSURE] Every narrow physical effect has a sufficient narrow physical cause.

[BROAD CLOSURE] Every broad physical effect has a sufficient broad physical cause.

According to interactionism, the mental is sometimes a (non-overdetermining) cause of the physical. But does the mental cause a fundamental physical property or a higher-level physical property? Here it is helpful to make a distinction between *Robust Interactionism* and *Modest Interactionism*. Whereas robust interactionism claims that the mental can (non-overdeterminingly) cause fundamental physical effects, modest interactionism asserts that the mental can *only* cause higher-level physical effects (in a non-overdetermining way). In the case of robust interactionism, neither NARROW CLOSURE nor BROAD CLOSURE is true; but in the case of modest interactionism, although BROAD CLOSURE is false, NARROW CLOSURE could still stand. Both robust interactionism and modest interactionism can satisfactorily

accommodate mental causation. This general interventionist framework of supervenient causation that I advocate leaves it open which version of interactionism is correct.⁵⁴

V. CONCLUSION

Contrary to most interventionists, I propose that a suitable intervention should hold fixed the variables in the context of supervenient causation. More specifically, we should hold the subvenient property fixed at *absence* while intervening on the supervenient property, and hold the supervenient property fixed at *presence* while intervening on the subvenient property. This approach promises to provide a proper interpretation of both the presence and absence conditions and to accommodate the proportionality constraint on causation. My proposal is a result of mutual coherence between two sets of theoretical considerations: general interventionist requirements on manipulation and fixation on the one end, and particular causal verdicts (such as causal judgments about the two pigeon examples) on the other end. It is important to note that my account of supervenient causation is neither too permissive nor too restrictive: while in some cases subvenient properties rather than supervenient properties should be regarded as the causes, in other cases causal efficacy should be attributed to supervenient properties instead of subvenient properties.

⁵⁴ In Zhong, “Sophisticated Exclusion and Sophisticated Causation,” *op. cit.*, I develop an *autonomy* approach to mental causation, which is a version of modest interactionism. In my most recent work, however, I argue that non-reductive physicalists should reject NARROW CLOSURE for solving the exclusion problem—that is, they should adopt robust interactionism. See Zhong, “Taking Emergentism Seriously,” *op. cit.*

REFERENCES

- Armstrong, D. M. (1978) *A Theory of Universals*, Cambridge: Cambridge University Press.
- Audi, P. (2013) "How to Rule out Disjunctive Properties," *Noûs* 47: 748-766.
- Baumgartner, M. (2010) "Interventionism and Epiphenomenalism," *Canadian Journal of Philosophy* 40: 359-383.
- Bennett, K. (2003) "Why the Exclusion Problem Seems Intractable and How, Just Maybe, to Tract It," *Noûs* 37: 471-497.
- Carey, B. (2011) "Overdetermination and the Exclusion Problem," *Australasian Journal of Philosophy* 89: 251-262.
- Dowe, P. (2000) *Physical Causation*, New York: Cambridge University Press.
- Endicott, R. (1993) "Species-Specific Properties and More Narrow Reductive Strategies," *Erkenntnis* 38: 303-321.
- Gillett, C. (2017) "Scientific Emergentism and Its Move beyond (Direct) Downward Causation," in M. P. Paoletti and F. Orilia (eds.), *Philosophical and Scientific Perspectives on Downward Causation*, New York: Routledge, 242-262.
- Hitchcock, C. (2001) "The Intransitivity of Causation Revealed in Equations and Graphs," *Journal of Philosophy* 98: 273-299.
- Hitchcock, C. (2001) "A Tale of Two Effects," *Philosophical Review* 110: 361-396.
- Horgan, T. (1993) "Nonreductive Materialism and the Explanatory Autonomy of Psychology," in S. Wagner and R. Warner (eds.), *Naturalism: A Critical Appraisal*, Notre Dame: University of Notre Dame Press, 295-320.
- Kim, J. (1998) *Mind in a Physical World*, Cambridge: MIT Press.
- Kim, J. (2005) *Physicalism, or Something Near Enough*, Princeton: Princeton University Press.
- Kolb, B. and I. Wishaw (2003) *Fundamentals of Human Neuropsychology*, 5th Edition, New York: Worth Publishers. 621-641.
- Kroedel, T. (2015) "A Simple Argument for Downward Causation," *Synthese* 192: 841-858.
- Lewis, D. (1973) "Causation," *Journal of Philosophy* 70: 556-567.
- Lewis, D. (1983) "New Work for a Theory of Universals," *Australasian Journal of Philosophy* 61: 343-377.
- Lewis, D. (2000) "Causation as Influence," *Journal of Philosophy* 97: 182-197.
- List, C. and P. Menzies (2009) "Nonreductive Physicalism and the Limits of the Exclusion Principle," *Journal of Philosophy* 106: 475-502.
- Loewer, B. (2007) "Mental Causation, or Something Near Enough", in B. P. McLaughlin and J. Cohen (eds.), *Contemporary Debates in Philosophy of Mind*, Oxford: Blackwell, 243-264.
- Mackie, J. (1974) *The Cement of the Universe*, Oxford: Oxford University Press.
- McDonnell, N. (2016) "Causal Exclusion and the Limits of Proportionality," *Philosophical Studies* 174: 1459-1474.
- Pearl, J. (2000) *Causality*, New York: Cambridge University Press.
- Raatikainen, P. (2010) "Causation, Exclusion, and the Special Sciences," *Erkenntnis* 73: 349-363.
- Robb, D. and J. Heil (2018) "Mental Causation," in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = < <https://plato.stanford.edu/archives/win2018/entries/mental-causation/>>.
- Salmon, W. (1984) *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.

- Schaffer, J. (2003) "Overdetermining Causes," *Philosophical Studies* 114: 23-45.
- Shapiro, L. (2010) "Lessons from Causal Exclusion," *Philosophy and Phenomenological Research* 81: 594-604.
- Shapiro, L. (2012) "Mental Manipulations and the Problem of Causal Exclusion," *Australasian Journal of Philosophy* 90: 507-524.
- Shoemaker, S. (2003) *Identity, Cause, and Mind*, Oxford: Oxford University Press.
- Shoemaker, S. (2007) *Physical Realization*, Oxford: Oxford University Press.
- Sider, T. (2003) "What's So Bad about Overdetermination?" *Philosophy and Phenomenological Research* 67: 719-726.
- Weslake, B. (2013) "Proportionality, Contrast and Explanation," *Australasian Journal of Philosophy* 91: 785-797.
- Wilson, J. (2011) "Non-reductive Realization and the Powers-based Subset Strategy," *The Monist* 94: 121-154.
- Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation*, New York: Oxford University Press.
- Woodward, J. (2008) "Mental Causation and Neural Mechanisms," in J. Hohwy and J. Kallestrup (eds.), *Being Reduced: New Essays on Reduction, Explanation, and Causation*, Oxford: Oxford University Press, 218-262.
- Woodward, J. (2015) "Interventionism and Causal Exclusion," *Philosophy and Phenomenological Research* 91: 303-347.
- Woodward, J. (2016) "Causation and Manipulability," in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/win2016/entries/causation-mani/>.
- Yablo, S. (1992) "Mental Causation," *Philosophical Review* 101: 245-280.
- Zhong, L. (2014) "Sophisticated Exclusion and Sophisticated Causation," *Journal of Philosophy* 111: 341-360.
- Zhong, L. (2017) "Semantic Normativity and Semantic Causality," *Philosophy and Phenomenological Research* 94: 626-645.
- Zhong, L. (2020) "Taking Emergentism Seriously," *Australasian Journal of Philosophy* 98: 31-46.