



McDonnell, N. (2017) Causal exclusion and the limits of proportionality. *Philosophical Studies*, 174(6), pp. 1459-1474. (doi:[10.1007/s11098-016-0767-3](https://doi.org/10.1007/s11098-016-0767-3))

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/130990/>

Deposited on: 21 November 2016

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Causal Exclusion and the Limits of Proportionality

Abstract

Causal exclusion arguments are taken to threaten the autonomy of the special sciences, and the causal efficacy of mental properties. A recent line of response to these arguments has appealed to "independently plausible" and "well grounded" theories of causation to rebut key premises. In this paper I consider two papers which proceed in this vein and show that they share a common feature: they both require causes to be proportional (in Yablo's sense) to their effects. I argue that this feature is a bug, and one that generalises: any attempt to rescue the autonomy of the special sciences, or the efficacy of the mental, from exclusion worries had better not look to proportionality for help.

Causal Exclusion and the Limits of Proportionality

1 Introduction

Causal exclusion arguments are taken to challenge the causal efficacy of higher-order properties such as blood type, air pressure, or acidity. According to such arguments, these higher-order properties cannot be considered genuine causes since the properties of fundamental physics – of quantum mechanics, perhaps – do all of the causal work without remainder. Since there remains no distinctive causal work left to do for the higher-order properties, they are not causes of malarial resistance, of wind speed, or of corrosion. This is a confronting conclusion for anyone who values the seemingly autonomous contribution of the special sciences, or for anyone who thinks that their mental properties – such as being angry, or in pain – have caused them to act.

There are myriad responses to such arguments but a new strategy has gained prominence in recent years. The strategy goes like this: instead of considering the plausibility of the specific causal claims within causal exclusion arguments individually, step back and consider those claims in light of an independently plausible theory of causation. Whilst such a strategy exists in earlier works,¹ the recent versions I have in mind adopt the so-called ‘Interventionist’ approach to causation.² In this discussion I will focus on two such contributions, the first from List and Menzies (2009, 2010), the second from Zhong (2014).

My contention is that the theories of causation on offer from List and Menzies and Zhong are not independently plausible at all. In fact, I will argue that these views

¹ See, for example, Yablo (1992).

² See for example: Hitchcock (2012), Raatikainen (2010), Shapiro (2012), Woodward (2015).

share a common flaw which can be traced to the very same feature that appeared to allow progress against causal exclusion arguments. I will argue that this observation has ramifications for any view that takes causation to be subject to a strong version of Yablo's *proportionality* constraint.

I will first briefly introduce the views of List and Menzies and Zhong in turn, highlighting what they have in common. I will then offer a critique of the causal theories on offer and show how the lesson generalises. I will conclude with some speculation about how this discussion should inform subsequent debates concerning exclusion arguments.

2 List and Menzies

In their 2009 paper "Non-Reductive Physicalism and the Limits of the Exclusion Principle", List and Menzies argue that adopting a *difference making* conception of causation allows a non-reductive physicalist to resist the master argument against them: Kim's causal exclusion argument. In particular, it allows them to resist the *Exclusion Principle* itself, which they characterise as follows:

Exclusion Principle: If a property F is causally sufficient for a property G , then no distinct property F^* that supervenes on F causes G . (p.478)

If we assume the causal closure of the physical world (as any physicalist must), and the supervenience of mental properties on distinct physical properties (as non-reductivists do), then this Exclusion Principle seems to rule out the causal efficacy of

the mental.³ Those unsatisfied with that conclusion incur the burden of identifying where this striking argument goes wrong.

List and Menzies propose the following conception of causation, broadly in the vein of Woodward's⁴ (2003): *F* is a cause of *G* iff *F* makes a difference to *G*. Making a difference is then analysed as follows:

Truth conditions for making a difference: The presence of *F* makes a difference to the presence of *G* in the actual world if and only if it is true in the actual world that (i) *F* is present $\Box \rightarrow G$ is present; (ii) *F* is absent $\Box \rightarrow G$ is absent. (2009, p482)

It is interesting to note how similar this conception of causation is to David Lewis's original counterfactual account from 1973. Lewis proposed two near-identical counterfactuals in his proposal for analysing causal dependence (the ancestral of causation itself):

Causal Dependence: event *e* causally depends on event *c* if and only if (i*) $O(c) \Box \rightarrow O(e)$; (ii*) $\sim O(c) \Box \rightarrow \sim O(e)$. (1973, p.563)

The only immediately obvious difference between the two is that Lewis's concerns the occurrence ('O') or non-occurrence ('~O') of events (*c*, *e*) and the List and Menzies conception concerns the presence or absence of 'variables' more generally

³ I will not go over the exclusion argument in detail here for brevity. Interested readers should consult the original List and Menzies and Zhong papers for the details and the caveats that I omit here.

⁴ The connection with Woodward is subtle in List and Menzies (2009), but explicit in (2010) p.108-109.

conceived. This is not a crucial difference for my purposes. What is crucial is that the semantics being employed for the counterfactual conditional differ between Lewis, and List and Menzies. In particular, List and Menzies adopt a weak centring, rather than strong centring, semantics for the counterfactual conditional. In simple terms, Lewis takes there to be just one world which is maximally similar to the actual world, *viz.* the actual world itself, whereas List and Menzies take it that, subject to the similarity ordering that the context invokes, it is possible that there are many worlds which are maximally similar to the actual world. In short, Lewis takes the first conditional to be trivial for any actually occurring events *c* and *e*, whereas List and Menzies take the first conditional to be substantial: it is a contingent matter whether (i) is true, even when *F* and *G* are both present in the actual world.⁵

This feature is critical to the List and Menzies view, as the following example will illustrate. Alice the pigeon is trained to peck all and only scarlet things presented to her. A scarlet patch is placed before Alice and she pecks. Now, by Lewis's account, the placing of a scarlet patch (*c*) and the placing of a red patch (*c**) both qualify as causes of the pecking. The scarlet patch was placed and Alice pecked (satisfying (i*)), and had the scarlet patch not been placed, Alice would not have pecked (satisfying (ii*)). Also, the red patch was placed and Alice pecked (satisfying (i*)), and had the red patch not been placed, Alice would not have pecked (satisfying (ii*)). According to List and Menzies however, the presence of red is not a cause of the pecking since, on their semantics, it is not enough that red was present and that the peck occurred for (i) to be true, it must also be the case that in all maximally similar worlds where the antecedent is satisfied (i.e. where the patch is red), the consequent is too (Alice pecks). We know, however, that if the patch that was placed were crimson,

⁵ Once again, I will not rehearse the details here. Interested readers are directed to their argument for this change in List and Menzies (2009) p.483–484.

and therefore still red, Alice would not have pecked. This means that when there are crimson patches in the vicinity (i) is false in the case of $F = \text{red patch}$, and so the placing of a red patch is not a cause, whereas the placing of a scarlet patch is to be considered a cause since both (i) and (ii) are true when $F = \text{scarlet patch}$.

This sort of distinction between the causal status of red and scarlet is familiar from Yablo (1992, p.273–279). Yablo argued that for F to be a cause of G , F had to be *proportional* with respect to G in the sense that F had to be specific enough to produce G , but not overly specific in respect of producing G .⁶ When $F = \text{red patch}$ in the pigeon example above, it is insufficiently specific with respect to Alice's pecking and so is *not* considered a proportional cause of that pecking. In the well-known alternative case of Sophie, who is trained to peck all and only red things, the results from the Alice case are reversed: $F = \text{red patch}$ is proportional to Sophie's pecking (it is just specific enough); whereas $F = \text{scarlet patch}$ is not since it misleadingly implies that a crimson patch would not have elicited a peck from Sophie (i.e. it is insufficiently specific). Importantly, the List and Menzies approach aligns with this notion of proportionality: when $F = \text{red patch}$ in the Sophie case, both (i) and (ii) are satisfied since in all cases where a red patch is placed Sophie will peck, and in all cases where nothing red is presented she won't; when $F = \text{scarlet patch}$ (i) is satisfied but (ii) is not, since if the scarlet patch had been absent, a crimson one could well have taken its place. So, List and Menzies and Yablo agree on this case. Moreover, List and Menzies take their view to independently confirm Yablo's proposal that causation is proportional, since they derive it from what they take to be an

⁶It is important to note that Yablo's analysis of this principle appealed to the determinate/determinable relation. In discussing this principle, I will mean only the claims about specificity of causes in respect of their effects, not Yablo's particular analysis of that principle.

independent, and independently plausible, conception of causation as difference making.

Returning to the Exclusion Principle, we can see that the case of Sophie provides a counterexample when the List and Menzies conception of causation is adopted. The presence of a red patch supervenes on the presence of a scarlet one, and the patch's being scarlet is *sufficient* for Sophie's pecking to occur, therefore, by the Exclusion Principle, the patch's being red cannot be a cause of the pecking. Yet, according to List and Menzies, the patch being red is exactly what caused the pecking. The conclusion drawn by List and Menzies is that those who proffer exclusion arguments rely on an inadequate conception of causation. Once this is resolved, so is the apparent problem with non-reductive physicalism. This is a striking result.

3 Zhong

A similar proposal has since been made by Lei Zhong in his 2014 paper "Sophisticated Exclusion and Sophisticated Causation." In this later proposal, the target is not Kim's causal exclusion argument, but rather an adapted version of it that Zhong considers a more sophisticated, and more resilient, problem for the non-reductive physicalist.⁷ Yet, whilst the target argument is slightly different, the overall strategy is the same: take a central premise in the argument and show that an independently plausible theory of causation is inconsistent with it. The target premise for Zhong is what he calls the *Causal Realisation Principle*:

⁷ See also Zhong (2011).

Causal Realisation Principle: If property A causes property B, then either : (I) A causes any supervenient property of B instantiated on this occasion (*upward causation*); or (II) A causes any subvenient property of B instantiated on this occasion (*downward causation*). (p.347)⁸

It is this principle that Zhong believes has been overlooked by previous versions of the exclusion argument, and it is this principle which he believes holds the key to resolving the exclusion problem.⁹

Invoking the work of Woodward, and the general success of *Interventionist* accounts of causation,¹⁰ Zhong proposes his *Dual Condition* conception of property causation (which he dubs ‘N’):

(N) A property F causes another property G if and only if

(N1) If an intervention that sets $F = f_p$ were to occur (while all other relevant variables in the causal graph are fixed), then $G = g_p$; and

(N2) If an intervention that sets $F = f_a$ were to occur (while all other relevant variables in the causal graph are fixed), then $G = g_a$. (2014, p.344)¹¹

⁸ I have substituted (i) for (I) and (ii) for (II) for clarity of reference.

⁹ Zhong argues for this in his (2011) and (2014) but I don’t discuss the details here since whether he is correct or not about the exclusion literature does not impact what I have to say.

¹⁰ See Zhong (2014) p. 344, and Woodward (2003) p.98 for details.

¹¹ I have exchanged Zhong’s X and Y for F and G respectively to help draw out the similarity with List and Menzies’ proposal.

In Zhong's usage, the upper-case letters pick out property types (variables) which can either be set to value *present* (lower-case letter with subscript *p*) or absent (with subscript *a*) and so he refers to N1 as the *presence condition* and N2 as the *absence condition*. It should be apparent that this very closely resembles the List and Menzies conditions for difference making: the first condition states that if F is present, G is present, and the second states that if F is absent, G is absent. If both are true, F caused G.

The similarities do not end there. Zhong also rejects the trivial reading of the presence condition when F and G are set to their actual state (present or absent). Echoing List and Menzies' rejection of strong centring, Zhong points out that this would fail to rule out insufficiently specific causes (such as $F=\text{red}$ in the Alice case above). This is the key feature of Zhong's conception of causation as it is what undermines the Causal Realisation Principle.

To illustrate, let us embellish the pigeon set-up. Suppose that the placing of a patch in front of either Sophie or Alice is a mechanical process triggered by the pressing of a button. Suppose further that the process only guarantees that some red patch will be placed, not which shade that patch will be (thus the outcome can be multiply realised).

In this case, let us take it as uncontroversial that the placing of the red patch (F) is caused by the pressing of the button (B), as this comes out true on all the accounts under consideration here. Now, the patch's being coloured (F^+) supervenes on it's being red (F), and its being red supervenes on its being scarlet (F^-). Since all agree that B caused F, the Causal Realisation Principle states that it must be the case that either (I) B causes F^+ ; or (II) B causes F^- . Zhong thinks both are false and so, by extension, the Causal Realisation Principle is too.

Regarding (II), the pressing of the button is insufficiently specific to be considered a cause of there being a scarlet patch placed in front of the pigeon. This is exactly analogous to the placing of the red patch in the case of Alice: it is just not specific enough to be considered a cause of her peck. Thanks to the substantial, rather than trivial, reading of the conditional in Zhong's presence condition, (II) will always be false when F is multiply realisable.

Regarding (I), if the button is not pressed, and so the absence condition is met for B, does that mean the absence condition will also be met for F^+ : does not-pressing the button mean that no coloured patch will be placed? In some cases, perhaps, but not all. If the process of placing patches carries on placing green and blue patches until the button is pressed, then even when there is no button pressing, there will still be a coloured patch. The point here is that it is not a matter of general principle that (I) is true, but rather a matter of contingent fact.

So, (II) is always false in cases where the F is multiply realisable, and (I) is true only in certain cases, so it is not always the case that one or other of them is true. This refutes the Causal Realisation Principle and therefore undermines the "sophisticated" version of the exclusion argument that Zhong has unearthed.

It will be important in what follows to highlight certain additional features of Zhong's proposal. First, Zhong is appealing to an interventionist semantics for the relevant counterfactuals, not the possible worlds semantics endorsed by List and Menzies. (This may lead to some divergent results, and I will discuss this issue further in §4.) Second, the causal graph with which Zhong is working contains a higher-order variable (red), and at least two lower-order variables which can realise it (scarlet, crimson). These variables are logically connected such that if any realiser of red is set to present in the graph, the red variable must be too, and if the red variable is set to

absent, every realiser variable in the graph must be too. This combination of variables is non-standard in interventionism¹² and it allows for cases of what List and Menzies call *downwards exclusion*: where a higher-order property causes of some effect but where no lower-order realiser of it does (2009, p.477). The example of Sophie is one such case: Sophie’s pecking is sensitive to the presence/absence of red, not to the presence/absence of scarlet (or crimson, or any other single realiser). Thus red downwardly excludes scarlet on Zhong’s Dual Condition account of causation.

Zhong does not explicitly discuss this implication, but clearly does not want to allow that higher-order properties can exclude lower-order properties as causes in the case of *physical* effects (2014, p.350).¹³ Alluding to List and Menzies, he complains that such a result “violates the Causal Completeness of Physics” (p. 349). There are two remarks to make about this issue before setting it aside. First, it is not clear that downwards exclusion in cases with a physical effect must violate the principle of causal closure. As Weslake (forthcoming) points out in relation to the List and Menzies proposal, the lower-order property may be excluded as a full *cause* but can nevertheless be *causally sufficient* for the effect (as scarlet is for the pecking). Therefore, if the causal closure principle can be adequately expressed in terms of causal sufficiency, then the cases of downwards exclusion need not conflict with that principle. Second, whilst Zhong would appear to think that downwards exclusion *should* be ruled out on his account when the effect is a physical effect, this result does not follow directly from the Dual Condition account of causation that he has provided, but rather stems from an independent commitment to the causal

¹² As Woodward (2015) says: “it is standard to assume, that it is possible to intervene on every variable represented in a graph and to set each such variable to each of its values independently of the values to which other variables are set.” (p.312, fn9)

¹³ I thank an anonymous referee for bringing this to my attention.

completeness of physics (all of Zhong's examples begin by presupposing causation between the lower-order physical properties).

In what follows I will set these issues aside. None of the problems which I wish to raise for Zhong (or for List and Menzies) hinge on a commitment to cases with fundamental physical effects.

4 The Relation to Woodward's Interventionism

Both List and Menzies and Zhong claim to refute a key premise of their chosen version of the causal exclusion argument by adopting an independently plausible account of causation. In Zhong's case, the theory adopted is explicitly Woodwardian, and it is introduced as a "well-developed" (2014, p.342), "respectable" (p.349) account of causation. List and Menzies appeal to their adaptation of Woodward's difference making conception, which they consider to be "well-grounded" (2009, p.476) and independently plausible (2010, p.108-109). Thus, in each case the credentials of the causal theory on offer are not explicitly defended, but are instead taken to be inherited from the plausibility of Woodward's version of the interventionist theory. In this section and the next, I will first challenge their inheritance claim and then go on to show that the gap between their views and Woodward's allows for a host of problematic cases for the new proposals.

The key area of agreement between List and Menzies and Zhong is in their assessment of cases where the cause variable is not optimally specific with respect to the effect. In the example discussed earlier, both proposals agreed that the placing of a coloured patch was not a cause of Sophie's peck. List and Menzies derived this assessment from their preferred weak-centring interpretation of the semantics of the

counterfactual conditional.¹⁴ Zhong, by contrast, considers the presence condition to be substantial so as to rule out insufficiently specific causes (and cites List & Menzies in the process (p.345, fn14)). Regardless of how they motivated the conclusion, they agree with one another that causes cannot be insufficiently specific with respect to their effects. In the jargon, we might say that the proportional cause (F=red) *upwardly excludes* the insufficiently specific cause (F=coloured) in this case. Thus each agree with the first half of Yablo's proportionality constraint.¹⁵

List and Menzies are explicit about their commitment to the other half too: causes cannot be overly specific in respect of their effects, as when the placing of the scarlet patch is taken to be a cause of Sophie's peck. Whilst Zhong does not explicitly endorse proportionality, the case he makes against the Causal Realisation Principle requires that he also adopt the second clause: that in cases where the cause variable is multiply realisable, causes cannot be overly specific in respect of their effects (2010, p.345). As discussed above, in such cases the proportional cause (F=red) *downwardly excludes* the overly specific cause (F=scarlet). Thus, List and Menzies and Zhong are both committed to the possibility of *upwards exclusion* and the possibility of *downwards exclusion*, and therefore both are committed to some version of the proportionality principle for causation (though for the latter, only in cases of multiple realisability).

However, nowhere in the original statement of the interventionist account of causation that List and Menzies and Zhong both appeal to, does Woodward commit to

¹⁴ See List and Menzies (2009) p.483–487. The reasons given for adopting weak centering are, on the face of it at least, independent from a desire to get the proportionality result.

¹⁵ Again, I am referring to the proportionality constraint as a principle, not to Yablo's determinate/determinable analysis of that principle.

the proportionality condition, or downwards exclusion.¹⁶ In that original work he is, at best, neutral. In subsequent work, however, Woodward explicitly disavows both of these key ideas. Here I quote selectively from a particularly lengthy footnote:

“[There are] several points at which the treatment of the exclusion argument that follows differs from the treatment in List and Menzies, 2009, even though they also take themselves to be defending an “interventionist” treatment of the exclusion problem. List and Menzies take it to be necessary condition for X to cause Y that X satisfy a proportionality requirement with respect to Y. They argue on the basis of this requirement that under the right conditions [...] there is “downward exclusion” [...] In contrast, I do not regard proportionality as a necessary condition for causation [...] Hence no “downward exclusion” is present.” (2015, p.305, fn1)

Thus, the view of causation that List and Menzies and Zhong are offering is not Woodward’s view at all. In fact, Woodward disavows the very feature of their view that is operative in addressing the problems of exclusion.

There are two further important departures from Woodward’s original theory to mention. As already discussed, List and Menzies adopt a possible-worlds semantics when assessing the relevant counterfactuals, whilst Woodward (and Zhong) adopt an interventionist semantics which does not require a specification of the similarity ranking of worlds. This will lead to divergent assessments of causal cases where the relevant alternatives featured in the interventionist’s graph or model outstrip those

¹⁶ Woodward does later (2010) consider how one might pick from among the causal connections, which are proportional, in the service of providing a special sort of causal explanation. This is a different topic, however, and Woodward’s proposal is convincingly criticized by Franklin-Hall (2016).

that occur in the closest possible worlds.¹⁷ For example, if the interventionist model contains the variable *pain*, and its possible realisers *c-fibre firing*, and *d-fibre firing*, then it could be the case that if the c-fibres had not fired then a d-fibre one would have instead (when pain remains set to *present* in the graph), and some aversive behaviour would still have occurred. Thus, c-fibre firing is *not* a cause of the aversive behaviour because that behaviour is insensitive to c-fibre's presence or absence. However, if the closest worlds in which the c-fibres do not fire are those in which no other realiser of pain replaces it, then List and Menzies will say that c-fibre firing *is* a cause on this occasion. This is what makes cases of exclusion context-dependent on the List and Menzies approach (pain's causal role is *realiser-sensitive* on this occasion, they would say), and represents a departure from standard interventionism.

Zhong also departs from *standard* interventionism (i.e. Woodward, 2003) in an important way. 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). Discussing just this sort of case, Woodward (2008, p.240) says that he sees "no reason" to accept that an alternate realiser should be taken to be present unless there is some specific backup mechanism which guarantees it.¹⁸ List and Menzies (2009, p.482) on the other hand appear to side with Zhong: the standard of similarity of

¹⁷ I owe this observation, and the subsequent example, to an anonymous referee.

¹⁸ This sort of complication arises when the causal graph includes logically connected variables, and it around this issue that Baumgartner's criticisms of Interventionism (as applied to exclusion problems) centres. See Baumgartner (2010). In more recent work, Woodward (2015) has responded by revising the definition of an intervention, but this work, and the controversial issues concerning logically connected variables, are not discussed by Zhong.

worlds may be such that the closest non-scarlet worlds are worlds where an alternative realiser of red is present. This suggests a gap between Woodward's much-vaunted (2003) theory and the proposals from List and Menzies, and Zhong. (It is important to note, however, that in his latest contribution to the debate, Woodward (2015) does allow the holding fixed of of-path variables in certain contexts.¹⁹)

Perhaps disagreeing with Woodward (2003) on such issues is no bad thing, and perhaps adopting a proportionality constraint will improve the original theory, despite Woodward's objections. I think it does not, but before I go on to show why, it is worth noting that the apron strings have now been cut: List and Menzies, and Zhong, give different results to Woodward on causal cases, and so can longer lay claim to Woodward's "well-grounded" and "respectable" theory of causation.²⁰ Their proposal needs to be considered anew.

5 Problems with Proportionality

What makes Woodward's interventionist account of causation plausible, well-grounded and respectable, is that it handles a wide range of canonical cases of

¹⁹ Woodward has argued that his original (2003) definition of an intervention should only apply in what he calls "standard" models (2015, p.318) – i.e. models in which the variables are suitably independent of one another. In non-standard models (such as those where the variables are logically related) a new definition of an intervention is required ("IV*"), one which allows for the fixing of certain off-path variables (p334). Woodward seems to treat this as a clarification of the original theory rather than an amendment, but as it allows holding fixed off-path variables, it may be the sort of treatment that Zhong has in mind. This is not discussed in Zhong (2014) however, and the viability of Woodward's revised/refined (2015) proposal deserves fresh consideration.

²⁰ As noted by an anonymous referee, *interventionism* as an approach to causation is independent of the commitment to proportionality (or any other specific commitment of Woodward's). My claim here is only about the gap between Woodward's (2003) version of interventionism, and those theories from List and Menzies, and Zhong, which claim that version's respectability.

causation within a single, elegant framework. Other theories, notably process theories²¹ and simple counterfactual accounts²², have been shown to struggle with certain sorts of example – absence causation, overdetermination, failures of transitivity in particular – and it is the fact that interventionism appears to offer an improvement in dealing with these cases that gives it its pre-eminent status.²³ What is agreed upon amongst theorists of causation is that delivering a theory which conforms with common sense, on clear cut cases, is of critical importance. I intend to show here that the List and Menzies, and Zhong, views both deliver unpalatable verdicts on such common sense cases and so do not deserve to be considered independently plausible in the manner that they claim, at least not without substantial additional argument. I will give two related arguments: one that suggests List and Menzies, Zhong, and defenders of a strong version of proportionality more generally, are committed to an error theory of ordinary causal claims, and another which demonstrates the difficulty that strong proportionality theorists face in relation to a central case from the causation literature.

5.1 Error Theory

Our ordinary causal claims are rarely proportional.²⁴ Take these three ordinary causal claims:

1. “Socrates’ drinking hemlock caused him to die.”

²¹ Such as advocated in Dowe (2000), and Salmon (1984).

²² See in particular Lewis (1973).

²³ Though, for important criticism of the project, see Hall (2007), and Strevens (2007).

²⁴ Similar worries are raised in Bontly (2005).

2. “Suzy’s throwing the rock caused the window to break.”
3. “The dog’s bark caused the baby to wake.”

In the first case, Socrates’ death will not be sensitive to the presence or absence of hemlock, since other poisons could have done the job instead. This means that the absence condition for List and Menzies, and for Zhong, will be false for *hemlock*, but both the presence and absence conditions will be true for *poison*. As such, hemlock is *not* a cause of Socrates’ death, and the first claim is false.

List and Menzies may complain that, given their closest-world semantics for the relevant conditions, this claim is only false if the alternate realisers of poison (arsenic, strychnine, etc.) would be present if hemlock were not. Notice, however, that the structure of this case is exactly that of the Sophie case: there is a higher-order variable (poison/red), a lower-order realiser (hemlock/scarlet), and an effect (death/peck). It was an important feature of their argument that Sophie’s peck was not sensitive to the absence of scarlet, precisely because alternate realisers of red could be present even if scarlet was not. The conclusion that an alternative realiser would replace scarlet in this case (or replace the neural realiser of an intention in the other case they discuss (2009, p.478-479)) is derived without embellishment of the case: whether or not a crimson alternative is modally close is *underdetermined* by the original example. Thus it seems that this first example is on a par with the examples given. What is more, as Weslake (forthcoming) points out, this underdetermination makes it implausible that our causal judgements are sensitive to those contextual details.

The remaining cases have the same structure. In example 2, the window breaking is sensitive to whether or not something sufficiently heavy was thrown – a

cricket ball, or brick perhaps – not whether or not that thing was a rock. Thus, the throwing of a *rock* is not a cause, and the second claim is false, according to List and Menzies, and Zhong.

In example 3, the dog needn't have barked for the baby to have woken, any loud noise, such as the slamming of a door, or the backfiring of a car, would have sufficed. Thus the waking is not sensitive to the dog's barking or not, but rather to the presence or absence of *any* sufficiently loud noise, and so List and Menzies, and Zhong, will take claim 3 to be false.

Each of these causal claims features a putative cause that can be considered overly specific in respect of the stated effect, and so out of proportion with that effect. If proportionality is a necessary constraint on causation, that is, if proportionality is interpreted as a *strong* constraint on which causal claims are to be considered true, then each of the ordinary causal claims above (and many more besides) will be considered literally false.

Some philosophers have suggested that proportionality may provide us a way of selecting from amongst the true causal claims, those which are optimally informative.²⁵ This *weak*²⁶ reading of the proportionality constraint would fit with the literal truth of the everyday examples listed above. However, it also fits with the literal truth of the claim that the scarlet patch caused Sophie to peck, in the case where List and Menzies and Zhong require that it did not. Recall that this is not an optional feature of the views: in order to avoid counting both higher and lower order properties as causes of subsequent behaviour, List and Menzies, and Zhong, require that one is a

²⁵ For example, see Bernstein (2014), Weslake (2013), and Woodward (2008).

²⁶ Elsewhere, Shapiro and Sober (2012) pick out two readings of proportionality as *pragmatic* and *semantic* respectively. Since those causal claims which violate the 'pragmatic' reading are clearly assertable, I find it more perspicacious to use the weak/strong labelling instead.

cause to the exclusion of the other. That is the core operative feature in each of their respective views and it is far from obvious that this could survive a shift to a weak interpretation of the proportionality principle.

Perhaps the best response List and Menzies, and Zhong, or any defender of the strong proportionality principle, could give would be to claim that the *genuine* causal relation relates only the proportional causes and their effects, but that pragmatic considerations may push us to over or under specify in particular cases. Thus, the claims above are literally false but they remain acceptable nevertheless. This amounts to an error-theory about ordinary causal claims.

No argument is offered to support such a confronting consequence, but that alone does not make it implausible. However, even if an error-theory about ordinary causal claims were to be plausible, that would not justify adopting an error-theory about well-considered, and widely accepted, causal claims from the philosophical literature. It is to such an example that I now turn.

5.2 Early Pre-emption

Here is a classic case of Early Pre-emption:

EP: Billy and Suzy are out to vandalise. Suzy picks up the only rock and throws it towards the window. The rock strikes and the window breaks. If Suzy had not thrown the rock, Billy would have (and he is a notoriously accurate shot).

In this case Suzy's throw caused the window to break, but the window's breaking did not depend on Suzy because Billy was there to step in. A very simple counterfactual account of causation, in which causation just is counterfactual dependence between distinct events, cannot be correct in light of this example. For this reason Lewis, in his original statement of his counterfactual theory, specified that causation is not identical with counterfactual dependence, but rather with *chains* of counterfactual dependence (1973, p.563). In EP above, there is a chain of such dependence that runs from Suzy (*c*) to the window (*e*) via some midpoint (*d*). Step 1: if Suzy had not thrown the rock ($\sim Oc$) then the rock would not have passed through that midpoint ($\sim Od$). Step 2: by the time the rock reaches the midpoint, Billy has been frustrated and has thrown no rock. At that point it is true (on the specified no-backtracking reading of the counterfactual) that if the rock had not passed through the mid-point ($\sim Od$) then the window would not have broken ($\sim Oe$). In step 1 we have a counterfactual dependence between *c* and *d*, and so *c* is a cause of *d*. In step 2 we have counterfactual dependence between *d* and *e*, and so *d* is a cause of *e*. By the transitivity of causation, *c* is therefore a cause *e* in the EP case according to Lewis's original account.²⁷ This is taken to be the correct result.

Woodward can also easily handle cases such as EP. In his original discussion of similar cases (2003, p77-78), Woodward argues that we should hold fixed Billy's lack of throw when considering whether or not Suzy's throw was a cause of the window's breaking. Since the window breaking is sensitive to Suzy's throw when we control for Billy, that throw is a cause of the window's breaking. Again, I take this to be the correct result.

²⁷ For the full details, see Lewis (1973) and the subsequent "Postscripts" (1986).

Before I turn to considering how List and Menzies, and Zhong, treat this case, note that I am using a case of *early* pre-emption, rather than so-called *late* and *trumping* varieties.²⁸ The latter sort are more difficult examples that could be considered everybody's problem, whereas the case I am using is readily handled by many accounts of causation (including Lewis's and Woodward's) and could reasonably be expected to be handled by any plausible account of causation. It is also a convenient case to use as the causation literature is univocal about what the correct causal ascription is in this case: Suzy is a cause of the window's breaking.

As with the simplest counterfactual account, the EP case presents a problem for List and Menzies, and Zhong. Due to the presence of Billy, the window breaking does not depend on Suzy, and so both List and Menzies' (ii) requirement, and Zhong's absence criterion, are not met for F=Suzy's throw. Thus, they cannot get the right result in this canonical case with the resources they have presented so far.

Arguably, both requirements are met by the more proportional claim that a *child's throw* caused the window to break: in all relevant cases where a child throws a rock, the window breaks, and when a child does not, the window does not break. So, List and Menzies and Zhong could argue that F=child's throw is what causes the window to break. Since the 'child's throw' could be realised by Suzy throwing the rock, or by Billy doing so, it is multiply realisable, and the window breaking is *realiser insensitive* (in List & Menzies' terminology) with respect to the throwing of the rock: either child will suffice. This means that both List and Menzies, and Zhong, are committed to the downwards exclusion of F=Suzy's throw by F=child's throw and

²⁸ See Lewis (1986), and Schaffer (2000), for the standard statements of late and trumping varieties respectively.

thus that ‘Suzy’s throwing the rock caused the window to break’ is false.²⁹ This result contradicts both common sense, and the consensus in the literature and yet is not addressed in either List and Menzies’ or Zhong’s work.

Perhaps List and Menzies can mimic Lewis’s solution instead. Take some midpoint that the rock passes through as a result of Suzy’s throwing it, but which it would not have passed through if Billy had instead. To say that the rock’s passing through that midpoint was caused by a child’s throw is to be insufficiently specific in respect of the effect, and thus is false. To say that the rock passing through that midpoint was caused by Suzy’s throw, however, is proportional. Let us also suppose that the rock’s passing through that point was also a proportional cause of the window breaking (since the relevant counterfactuals must still be read in a non-backtracking way). Thus, by the transitivity of causation they could claim that Suzy’s throw is a cause of the window breaking. This delivers the consensus result for EP. However, whilst each step in the causal chain was proportional, the first link (Suzy’s throw) is not proportional with respect to the last (the window breaking). Thus, EP demonstrates that *proportional* causation is not transitive.³⁰

Zhong, on the other hand, may instead look to replicate Woodward’s treatment of the case: insist that Billy’s failure to throw should be held fixed when considering

²⁹ Notice that, if correct, this would represent a striking result. It would, at a stroke, resolve all cases of pre-emption! On this view, we should have said all along that Suzy’s throw was not a cause of the window breaking, but that a child’s throw was. This would turn 40 years of the literature on its head. But it is precisely that this is not a plausible treatment of pre-emption cases that means it has not been adopted.

³⁰ An anonymous referee offers a friendly suggestion for the defender of proportionality: mimicking Lewis, take causation not to be identical with proportional counterfactual dependence, but instead chains thereof. I think this is an interesting suggestion, one I have pursued independently (McDonnell, *Transitivity and Proportion in Causation*, manuscript), however as it allows that causes and their effects at either end of a proportional chain can be out of proportion with each other, it will only help the defender of a *weak* proportionality constraint, not any of my targets in this paper.

Suzy's causal status.³¹ However, given the close parallel with the Sophie case, Zhong cannot consistently adopt this line. In the pigeon example, Zhong did not hold fixed the status of the alternative realiser (crimson) when considering what would occur in the absence of scarlet, and it was this that allowed him to conclude that the effect behaviour was not sensitive to the presence or absence of scarlet. I pointed out in §4 that this was a non-standard approach for interventionists, who typically assume that the variables in the model will not be logically connected. In the EP case, then, Zhong cannot consistently claim that Billy's throw – which is an alternative realiser for the child's throw variable – be held fixed.

So, the view of causation that List and Menzies, and Zhong, adopt commits them to an error-theory about ordinary causal talk, to the denial of the transitivity of causation, and to the denial of a canonical causal truth in the EP case. There may be good arguments to be made in favour of these commitments – they would not be alone in denying transitivity, for example³² – but absent those arguments List and Menzies, and Zhong, cannot claim to be offering a “well-founded”, “respectable” and independently plausible account of causation. This was a central claim in each of their treatments of the exclusion problem.

Furthermore, the lesson generalises. Any view which adopts a strong, or semantic, treatment of the proportionality principle owes some argument in light of these three features. In particular, given that the standard methodology in the philosophy of causation is to progress by appeal to ordinary examples, anyone

³¹ An anonymous referee points out that List and Menzies may also be able to give something like this response by insisting that the right counterfactual in these cases is one with a *complex* antecedent: if Suzy hadn't thrown *and* Billy still didn't throw, then... Such a proposal would need to be clearly laid out before it could be assessed, but the argument I give here against Zhong still applies: what justifies the different treatment of EP and the Sophie case?

³² See, for example, Hitchcock, (2001), and McDermott, (1995).

offering such a theory must explain how they can come to the conclusion that ordinary causal claims are systematically false, without cutting off the methodological branch they used to get there.

6 Conclusion

I have argued here that the proportionality principle cannot be a strong constraint on causation, as List and Menzies, and Zhong, require. Here I offer some brief speculative remarks about what it might constrain.

Borrowing from Weslake (2013), I think we should take proportionality to be a ‘dimension of explanatory value’. I take this to mean that in general, though defeasibly, a more proportional causal claim will provide a better explanation. This has the merit of capturing the sense in which more proportional causes seem preferable, without committing us to an error theory of out-of-proportion causal claims. Thus, I take proportionality to be a defeasible constraint on optimal causal explanations.

Recast this way, the appeal to proportionality is not fit to resist exclusion arguments, since those arguments concern causation – the *causal* closure of the physical, the denial of *causal* overdetermination – not causal explanations. An analysis of the difference between causation and causal explanations is a topic for another paper, but granting that the two do come apart (as my preferred reading of proportionality assumes), then List and Menzies, and Zhong, would appear to have been aiming at the wrong target.

Or perhaps not. The urge to resist exclusion arguments stems from the urge to resist reductionism: the reduction of all special sciences to fundamental physics. In

their discussion, List and Menzies (2010, p.122-124) draw close parallels between their work, and that of Alan Garfinkel (1981) who argues that higher-order *explanations*, as provided by the special sciences, are indispensable, and thus resists *explanatory* reduction. I contend that it is entirely possible that the causal truths reduce to those found in the fundamental physics, even if our casual explanations do not. Here is Strevens proposing something very similar in the preface to his book on explanation:

“I do not assert that biological phenomena can be explained in the same way as physical phenomena. I assert only that the raw causal ingredient is the same in biology as in physics...” (2008, p. iii)

This sort of view would vindicate the anti-reductionist urge read as concerning causal explanations, and make sense of the force and influence of exclusion arguments, which concern causation itself.

I conclude, then, as follows. Those who wish to defend the autonomous contribution the special sciences make to our causal knowledge, or the efficacy of our mental lives, should not look to proportionality as a response to causal exclusion arguments. Instead, they should question whether or not causal exclusion arguments really threaten what it is that they care about.

References

Baumgartner, M. (2010). Interventionism and Epiphenomenalism. *Canadian Journal of Philosophy* 40 (3):359-383.

- Bernstein, S. (2014). Two Problems for Proportionality about Omissions. *Dialectica* 68 (3): 429-441
- Bontly, T.D. (2005). “Proportionality, causation, and exclusion”. *Philosophia* 32 (1-4): 331-348
- Dowe, P. (2000). *Physical Causation*. Cambridge University Press
- Franklin-Hall, L. R. (2016). High-Level Explanation and the Interventionist's 'Variables Problem'. *British Journal for the Philosophy of Science* 67 (2): 553-577
- Garfinkel, A. (1981) *Forms of Explanation*. New Haven: Yale University Press.
- Hall, N. (2007). “Structural equations and causation”. *Philosophical Studies* 132 (1): 109 – 136
- Hitchcock, (2001), “The intransitivity of causation revealed in equations and graphs”, *Journal of Philosophy* 98 (6): 273-299
- Hitchcock (2012). “Theories of Causation and the Causal Exclusion Argument”. *Journal of Consciousness Studies* 19 (5-6)
- Lewis, D. (1973). “Causation”, *Journal of Philosophy* 70 (17): 556-567
- Lewis, D. (1986). (1986) “Postscripts to ‘Causation’”, in *Philosophical Papers Vol. II*. Oxford University Press
- List, C. and Menzies, P. (2009). “Non-Reductive Physicalism and the Limits of the Exclusion Principle”, *The Journal of Philosophy* 106, 9 (September 2009): 475–502
- McDermott, M. (1995). “Redundant causation.” *British Journal for the Philosophy of Science* 46 (4): 523-544.

- Menzies, P. and List, C. (2010). "The Causal Autonomy of the Special Sciences". In Cynthia McDonald and Graham McDonald (eds.), *Emergence in Mind*. Oxford University Press.
- Raatikainen, P. (2010). "Causation, exclusion, and the special sciences". *Erkenntnis* 73 (3): 349–363,
- Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.
- Schaffer, J. (2000), Trumping preemption. *Journal of Philosophy* 97 (4): 165-181
- Shapiro, L.A. (2012). "Mental Manipulations and the Problem of Causal Exclusion". *Australasian Journal of Philosophy* 90 (3): 507–524
- Shapiro, L.A. and Sober, E. (2012). "Against proportionality", *Analysis* 72 (1): 89-93.
- Strevens, M. (2007). "Review of Woodward, *Making Things Happen*". *Philosophy and Phenomenological Research* 74 (1): 233–249.
- Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press.
- Weslake, B. (2013). "Proportionality, contrast and explanation". *Australasian Journal of Philosophy* 91 (4): 785-797
- Weslake, B. (forthcoming). "Difference-Making, Closure and Exclusion." In H. Beebe, C. Hitchcock & H. Price (eds.), *Making a Difference*. Oxford University Press
- Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation* Oxford University Press: New York
- Woodward, J. (2008). "Mental causation and neural mechanisms". In Jakob Hohwy and Jesper Kallestrup (eds.), *Being Reduced: New Essays on Reduction, Explanation, and Causation*. Oxford University Press.

- Woodward, J. (2010). "Causation in biology: Stability, specificity, and the choice of levels of explanation." *Biology and Philosophy* 25 (3): 287-318
- Woodward, J. (2015), "Interventionism and Causal Exclusion", *Philosophy and Phenomenological Research*, 91, 2
- Yablo, S. (1992). "Mental Causation", *Philosophical Review* 101 (2): 245-280
- Zhong, L. (2011). "Can Counterfactuals Solve the Exclusion Problem?" *Philosophy and Phenomenological Research*, vol. 83, p.129–147
- Zhong, L. (2014). "Sophisticated Exclusion and Sophisticated Causation", *The Journal of Philosophy* 111, 7 (July 2014): 341–360