McDonnell, N. (2019) The non-occurrence of events. *Philosophy and Phenomenological Research*, 99(2), pp. 269285.

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.

This is the peer reviewed version of the following article:

http://eprints.gla.ac.uk/149496/

Deposited on: 10 October 2017

THE NON-OCCURRENCE OF EVENTS

Neil McDonnell

University of Glasgow

Introduction

Counterfactual treatments of causation hinge, one way or another, on whether the non-occurrence of one actual event would have led to the non-occurrence of another. We might reasonably ask: what does it take for an event not to occur? This question has been significantly under-explored. In what follows I will introduce what I call *the non-occurrence problem*, and argue that *if* one wishes to adopt a simple counterfactual analysis of causation, then the standards of non-occurrence for the cause and the effect should differ. I will then attempt to render the non-occurrence standards more precise, and show that the double standard offers novel insight into the question of the transitivity of causation. I will then show that the issue that underpins the non-occurrence problem is not limited to simple counterfactual accounts, but also applies in a different form to more recent extensions to the difference making family: contrastivist and interventionist accounts of causation.

The Non-occurrence Problem

How much delay or change do we think it takes to replace an event by an altogether different event, and not just by a different version of the same event? An urgent question, if we want to analyze causation in terms of the dependence of whether one event occurs on whether another

event occurs. Yet once we attend to the question, we surely see that it has no determinate answer.

(Lewis 2004, 186)

Lewis's original (1973) counterfactual analysis of causation states that one event is a cause of another iff there exists a chain of counterfactual dependence between them. One event e counterfactually depends on another c iff the following counterfactual conditional (or causal conditional) is true:

$$\neg Oc \ \square\!\!\rightarrow \neg Oe$$

This is to be read as follows: if event c had not occurred, event e would not have occurred.[1] This intuitively works well in a range of simple cases:

1. If Socrates had not drunk the hemlock, he wouldn't have died. So, Socrates' drinking hemlock caused him to die.

2. If Striker hadn't taken the shot, Keeper wouldn't have dived. So, Striker's shot caused Keeper to dive.

3. If Michael Jackson hadn't died, Cristiano Ronaldo would still have signed for Real Madrid. So, Michael Jackson's death did not cause Cristiano Ronaldo's transfer.[2]

Consider, though, *some* of the ways that antecedent might be realised. For example, if Socrates had thrown himself from a cliff instead of drinking hemlock, if Striker had thrown a grenade instead

---

[1] I am glossing some details here: the events must be distinct, and the counterfactual must be read in a special 'non-backtracking' way (see Lewis, 1979). My discussion in this paper is about this causal counterfactual – the conditional that simple counterfactual theories of causation require – it is not about counterfactual conditionals in general.

[2] Background: both of these events took place on June 25th 2009. For illustration I will be assuming Jackson died before Ronaldo signed.

of taking the shot, or if Jackson had sabotaged Ronaldo's deal instead of taking pain medication. *If* these were the right ways to read the antecedent of each counterfactual, then the consequents—and thus the counterfactual theory's causal verdict—would quite plausibly have been reversed. Socrates would still have died, so drinking hemlock didn't cause the death. Keeper would still have dived out of the way, so Striker's shot would not have caused the actual dive. Ronaldo wouldn't have signed for Real Madrid, so Jackson's actual death *was* a cause of the transfer.

According to Lewis, these are not the right ways to read the counterfactual, precisely because if we simply look to *any* world that satisfies the antecedent, then all bets are off as to what that world will be like in other respects, particularly in respect of the effect event we are interested in. For this reason, we are to consider only the closest antecedent worlds, that is: the worlds which deviate least from actuality in order to satisfy the antecedent. We should not, for example alter the location of Socrates to the top of some tower, or alter the outcome of his trial. We should not imagine that the Striker had a grenade, or that Jackson was part of an anti-Ronaldo conspiracy. Instead, we should hold the facts of actuality fixed as far as possible, and consider only a minimal alteration to make the antecedent true.[3] [4]

Notice, though, that just as extravagant ways to realise the antecedent turned causes into non-causes (in the cases of Socrates and Striker), overly conservative deviations from actuality can too. In

---

[3] I note here that it is a vexed question how we should rank the closeness of worlds. Lewis gave a ranking schema in his (1979), but it has been widely criticised (see, *inter alia*, Krasner and Heller 1994; Bennett 2003; Schaffer 2004; Hawthorne 2005; Kment 2006; Hájek manuscript) and I think there is good reason to think that some alternative account is required. However, for the purposes of this paper, I will simply follow Lewis's schema since it is well known, and it will help focus on my target issue regarding events. I will briefly consider the issue again in §3.3.

[4] Carolina Sartorio (2010) argues (note 7, p 266) for a reading of the counterfactuals such that a specific alternative event should replace the actual event in cases where we know what an agent's alternative preference might be. I disagree, but my case here is not built on that disagreement. See notes 10 and 13 below..

the very closest worlds where the actual event of Socrates' drinking hemlock does not occur, we might think that a very close alteration of that event will take place instead. But a very close alteration, will bring about a very similar outcome: if Socrates had drunk just a little less hemlock (and thus the actual drinking event failed to occur) he still would have died. Combining the counterfactual account of causation, with this reading of the antecedent, renders the result that Socrates' drinking hemlock was *not* a cause of his death. That is wrong. Here is Lewis pointing out the problem, and glossing a solution:

> What is the closest way to actuality for C not to occur? It is for C to be replaced by a very similar event, one that is almost but not quite C, one that is just barely over the border between versions of C itself and its nearest alternatives. But if C is taken to be fairly fragile, then, if almost-C occurred instead of C, very likely the effects of almost-C would be almost the same as the effects of C. So our causal counterfactual will not mean what we thought it meant, and it may well not have the truth value we thought it had. When asked to suppose counterfactually that C does not occur, we do not really look for the very closest possible world where Cs conditions of occurrence are not quite satisfied. Rather, we imagine that C is completely and cleanly excised from history, leaving behind no fragment or approximation of itself. (Lewis 2004, 90)

This *excision* standard, as I will call it, is intended rule out both the extravagant replacements which I initially considered, and the overly conservative ones, which we get on the rival *change* standard, in one go. It is supposed to be a solution to the problem of specifying the conditions for an event's non-occurrence. However, nowhere are we told what *complete excision* amounts to, and it is past time that we sought to understand and critically assess this solution to the non-occurrence

problem.[5] That is the aim of this paper. There is no further ambition to defend or endorse counterfactual analyses of causation, there is only the conditionalised aim of establishing what the best way of dealing with the non-occurrence problem is, *if* one is to adopt a simple counterfactual approach to causation.

I will proceed as follows: I will first consider the two rival standards for non-occurrence (excision and change) in §2 and argue that one standard applies to the cause, and the other to the effect. I will then consider how we might define those standards in §3. In §4 I consider how this treatment might have been overlooked to date, how it might shed light on certain counterexamples to the transitivity of causation, and how all this relates to the issue of causal *proportionality*. In §5 I relate these issues to contrastive and interventionist causal theories, and conclude.

Excision versus Change

Lewis distinguished two standards of non-occurrence that we might consider, the *excision* standard (EX), and the *change* standard (CH) as I will call them. Here is a case that shows the two standards coming apart. Suppose that Suzy has thrown a heavy red ball at a wine bottle, and that the bottle has broken on contact. We might say:

1. If Suzy hadn't thrown the heavy, red, ball, the wine bottle wouldn't have broken.

If we apply the EX standard to Suzy's event, and consider the closest worlds in which the throw was *completely and cleanly excised from history leaving no fragment or approximation of itself*, then the

---

[5] It is clear from what Lewis goes on to say after the quoted passage that the non-occurrence issue was an important part of the motivation behind the 'Influence' account of causation he introduced in that paper. Lewis's (2004) proposal is not my focus here and so partly in the interests of clarity, and partly to avoid unnecessary exegetical controversy, I avoid detailed discussion of it in this paper. That said, I highlight some relevant parallels via the footnotes in section 5.1.

counterfactual comes out as true, just as the counterfactualist about causation might hope. However, if we apply the CH standard, we might suppose that the throwing of the heavy red ball is replaced by the throwing of a heavy green ball instead. On the assumption that the colour of the projectile had nothing to do with the bottle breaking,[6] this will render the counterfactual false.

The supposition that the hue is what might be what is altered in the closest alternate case is purely illustrative since it is entirely plausible that there are closer alterations where the ball is but a fraction of a gram lighter, or merely has a pair of atoms permuted. In each case, the window breaks even when the putative cause is so *changed*. This example demonstrates why Lewis suggested an EX standard for the non-occurrence of events within counterfactual theories of causation.

Notice, though, that the two standards under consideration here could in principle apply differently to the cause and to the effect. There are four possible combinations: Both cause and effect can be held to excision standards (EX-EX); both cause and effect can be held to change standards (CH-CH); cause can be held to an excision standard and effect to a change standard (EX-CH); or vice versa (CH-EX).

As a first gloss, let us take it that the complete excision of an event requires that it is completely altered such that no part or aspect of that event remains (i.e. leaves no *fragment or approximation of itself*), and let us further take it that a changed event is one that differs in *any way* (i.e. *just barely over the border between versions of C itself and its nearest alternatives*). Thus we can paraphrase our four possible alternative treatments as follows:

---

[6] This is a conceit. Our physics is such that were we to make whatever alterations required to the ball to change its colour, *some* difference would be made in the bottle: the mass and charge of the surface of the ball would affect the bottle in some minute way. I'll pretend that our physics is not this way for the purposes of illustration only.

c is a cause of *e* if and only if:

EX-EX completely altering *c*, completely alters *e*.

CH-CH altering *c* in any way, alters *e* in some way.

CH-EX altering *c* in any way, completely alters *e*.

EX-CH completely altering *c*, alters *e* in some way.

It will be useful to test these alternatives in light of the above example of Suzy's throwing the heavy red ball at the wine bottle. Clearly, Suzy's throw caused the bottle to smash. A causal analysis must agree with this claim to be acceptable. Now, consider the verdicts that the counterfactual analysis would give on each of the four different standards of non-occurrence:

EX-EX will be too strong a requirement as even the complete excision of Suzy's throw would leave some aspect of the bottle smash unaltered: the colour of the glass for example, or the shape of the base. EX-EX gets the test case wrong.

CH-CH is also too strong a claim. We are supposing that altering the hue of the ball does not alter the bottle breaking event in any way, so it is not true that *any* change in *c* will bring about some change in *e*. CH-CH gets the test case wrong too.

CH-EX is an even stronger claim than CH-CH. If altering the hue doesn't alter any aspect of the bottle break, then *a fortiori* it cannot alter *e* completely. CH-EX gets the test case wrong.

EX-CH is a much weaker claim. EX-CH is satisfied if the complete excision of the cause, i.e. Suzy's throw, makes any difference to the effect, i.e. the bottle breaking. EX-CH is the only standard that gets the test case right since the complete removal of Suzy and the ball from the environment result in the bottle remaining intact.

This is a surprising result. Neither of the symmetrical standards (EX-EX or CH-CH) under consideration get the test case right, therefore this Suzy case stands as a counterexample to any theory that proposes a uniform CH or EX standard of non-occurrence for both the cause and the effect in a counterfactual theory. Thus, this example stands as evidence for my central claim in this paper: *if* one wishes to adopt a simple counterfactual analysis of causation, then the standards of non-occurrence for the cause and the effect should differ.

Note that this is a simple test of the available combinations once we realise that there is an ambiguity in the notion of a non-occurrence of an event. The test case used is not an extraordinary or complex case, or one that generates known problems so, pending argument to the contrary, I think that it is reasonable to take the results at face value. The lesson here seems to be that *if* one wishes to adopt a simple counterfactual analysis of causation, then the original analysis needs to be amended to reflect a double standard in the treatment of what it is for the cause and effect not to occur: we must look to the complete excision of the cause, and consider if the effect is thereby changed.

In this section I have offered only *prima facie* evidence of this asymmetry by way of an example, but in what follows I will refine the proposal and consider issues around the semantics, how we might have failed to notice the asymmetry before, and how the existence of the asymmetry would explain some of the problem cases around the transitivity of causation. The remainder of this paper does not merely trace the implications of the asymmetry claim made here, but rather it bolsters the case for that claim by demonstrating its explanatory force.

Refining EX-CH

The foregoing makes the case for amending the the original counterfactual test for causation. In this section I consider a problem for the initial gloss of EX and CH given above, and use that problem to motivate a contextualist refinement. I will then seek to incorporate the refined definitions into the semantics of a more plausible counterfactual analysis of causation.

Too Many Causes

I pointed out above that the EX-CH standard for non-occurrence was the weakest under consideration, and it turns out that it will prove much too weak to fit with our ordinary causal discourse. Given our physics, the complete excision of any event in the backwards light cone of some region, will alter *some* part or aspect of that region. This is due to the minute gravitational influence that even tiny quantities of matter exert. So, given our physics, the EX-CH standard will consider every event in the backwards light cone of e, a cause of e.

Lewis spotted this issue in another context, in his consideration of so-called *fragility* responses to late pre-emption problems. In traditional late pre-emption cases, there is a back-up event which guarantees that the effect will occur, albeit a moment later, even in the absence of the actual cause. This undermines the dependence of the effect upon the cause, and confronts dependence-based analyses of causation in general. The fragility response to pre-emption adopts a highly sensitive (i.e. fragile) modality for the effect event. This distinguishes the effect as it actually occurs, from the very slightly different way it would have if caused by the back-up, and thus re-establishes the dependence of the actual event on the actual cause in pre-emption cases. Lewis dismissed this response since he thought that any theory which delivered the result that a dog's

barking several streets away could be a cause of a bottle's breaking (which he took the fragilicist approach to be committed to) thereby completed a *reductio* of that theory.[7]

Since the initial gloss I gave of the EX-CH standard for non-occurrence will render every event which alters the effect in any way, a cause of that effect, this standard of non-occurrence as it stands will be subject to the same 'too many causes' complaint that Lewis outlined against the fragilicist. The problem seems to be that in our ordinary judgements, not all features of the event are equal: we care about the *breaking* of the bottle, not its colour, or the shape of its base. What is relevant is that the bullet struck the victim, not the markings on the the metal, or the gravitational influence that some distant star exerted on it as it did so. So, in order to track such ordinary judgements, we will need to distinguish between the features of an event that are relevant in a given context, and those which are not. I will refer to such features as *essential* and *accidental* respectively, though in my usage this is intended to be a fairly neutral characterisation: no particular view of events, or metaphysically loaded conception of essence should be required for the purposes of this paper. The essential aspects of an event are those which characterise that event in a given context.[8] The initial gloss of the CH standard for non-occurrence treated an event as non-occurring if it was changed in at least one respect. The 'too many causes' problem showed that this was too weak a standard for non-occurrence, as it rendered every event in the backwards light cone of an event a cause of that event, on a counterfactual account. In order to ensure the relevance of the difference made to the effect event, I propose that we instead consider an event to have been sufficiently changed so as not to have occurred by the CH standard iff at least one essential feature is absent. Thus, when the relevant issue is that the bullet struck the senator, then a distant star, or colliding photons, will not be

---

[7] It is worth pointing out that a contextualist fragilicist need not accept this commitment in relation to the causal status of the dog's barking. See McDonnell (2016) for discussion.

[8] Though, see McDonnell (2016) for an opinionated stance on events and event modality.

considered a cause as they do not make a difference to the feature of the event we take to be

essential: that the bullet struck the senator. Similarly, the ball's striking the bottle is a cause of its

smashing, but the dog's barking is not, precisely because the ball made a difference to the feature of

the effect that we take to be essential (the smashing) but the dog's barking did not.[9][10]

3.2   The Void

When considering Lewis's excision standard for non-occurrence, Hall and Paul ask the following:

> What exactly does such "complete and clean excision" consist in? Removal of the event
>
> using some sort of metaphysical scalpel? Leaving behind... what? The Void? (Paul and
>
> Hall 2013, 51)[11]

The initial gloss I gave of EX stated that an event had been *excised* iff it had been completely

altered such that no part or aspect of that event remained, so the EX standard does indeed seem to

leave us with a void (unless the event we started with was one, of course, but that is another story[12]).

The sudden appearance of a void will have an enormous impact on every surrounding region, and

---

[9] I note here that some will take this as evidence that events should be individuated by their essential properties—(Lewis 1986) and (Kim 1973) most obviously—or that the causal relata are event aspects (Paul 2000). Whilst I disagree (see above reference), the issue is orthogonal here.

[10] The 'too many causes' problem I discuss here is a different problem from that of 'unwanted positive causes' identified by Carolina Sartorio (2010). Sartorio's problem arises around the issue of allowing omissions to be causes – a topic I studiously avoid in this paper – however, as noted in fn 13 below, some of the mechanics at play in Sartorio's approach are relevant to my discussion of how we read the causal counterfactual.

[11] I should note that Hall and Paul's discussion immediately moves on to a wholly different approach which requires distinguishing default and deviant states. This is a rival treatment of the non-occurrence problem, but one that I will not address here. The purpose of this paper is to give as charitable an interpretation as possible of the *complete excision* approach. Comparisons with rivals will only be appropriate once the complete excision proposal is suitably clarified.

[12] Absence causation is a topic that deserves its own discussion.

thus the complete excision of any (ordinary) event, will perturb all of its surroundings. Thus, by the lights of a counterfactual analysis, every event would be a cause of those surroundings. This would seem to count many more effects than common sense would endorse, and so presents the EX standard for non-occurrence, with a 'too many effects' problem to mirror the 'too many causes' problem above. Replacing every cause with a void, and seeing what results, is simply too blunt a test for causation by the lights of our common sense judgements.

As with the CH standard, perhaps the distinction between essential and accidental features can help. Let us suppose that completely excising the event does not require the removal of every feature (and thus creating a void), but rather the removal of every *essential* feature. In the Suzy example, the throwing of the ball seems essential to the cause event under consideration, whereas that particular arrangement of mass, or what Suzy was wearing, does not. So, by the updated standard for complete excision, we should consider the closest worlds where there is no throwing of the ball, not the closest worlds where a void replaces Suzy and the ball. What *exactly* Suzy does in those closest worlds will depend on what the smallest alteration from actuality is that will result in a non-throw, and this will vary with the context, however worlds where there is some mass in the region will be more alike actuality, than worlds where a void appears. The updated conception of *excision* avoids the 'too many effects' problem by excising only the essential properties of the event.[13]

---

[13] In her paper "The Prince of Wales Problem for Counterfactual Theories of Causation", Sartorio raises a series of problems for the counterfactual criterion (the claim that counterfactual dependence is sufficient for causal connection) based on a scenario in which the Prince of Wales fails to water some plants. It is important for Sartorio's argument (p264 – 266) that if the Prince had not done X, he would have done some specific Y instead, and so Sartorio's reading of the counterfactual applies a CH standard to the antecedent. This makes her's a rival view of how to read the counterfactual. Sartorio defendst her somewhat controversial reading against the charge that it requires so-called 'backtracking' (note 7, 275), but my argument above stands as a counterexample to any reading that applies a CH standard to the antecedent, including Sartorio's.

Given the similarity in the issues raised, it is worth considering whether the EX and CH standards are in fact just another way of picking out robust and fragile treatments of events respectively. Of course, what an excision or change of some event amounts to is highly dependent on the modality of that event, so the fragility or robustness *interacts* with the EX and CH standards when yielding a verdict. Nevertheless, I think that the EX/CH and fragile/robust distinctions are quite different, for at least two reasons. First of all, fragility and robustness are *relative* assessments of the modal status of the event: fragile events take relatively little alteration compared to others to be destroyed, robust events take relatively more. The distinction between the EX and CH standards are not *relative* to any events—they represent fixed standards for non-occurrence once the event's modality has been established. Secondly, the distinction between EX and CH standards for non-occurrence cross-cuts the fragile-robust distinction: we can consider the complete excision of a fragile event (no window smash takes place), or we can consider a near alternate instead (a very similar smashing, a moment later). A robust event may be robust in virtue of its tolerance of delay— the window breaking that happens a minute later is still the same event—but that event could be replaced by a slightly different alternative (where the window breaks differently) or completely excised (where the window doesn't break at all). Thus, the standards of non-occurrence considered here, and the fragility/robustness distinction, come apart. The non-occurrence problem is a problem in its own right.

Reflecting the Double Standard in the Semantics

Recall that the conditional at the heart of simple counterfactual analyses of causation (the causal counterfactual) has the following form:

$\neg Oc \, \Box\!\!\rightarrow \neg Oe$

If there is to be a different standard for the cause as for the effect, that difference will need to be reflected somehow in the semantics. There are three potential locations: (i) in the $\Box\!\!\rightarrow$ conditional, (ii) in the specification of the events themselves (c, e), or (iii) in the $\neg O$ clause.

The counterfactual conditional need not relate events, let alone their negation, so somehow encoding the different non-occurrence standards within the conditional (i.e. option (i)) would be unmotivated. Of course, the standard semantics for that conditional already encodes an asymmetry of sorts – the truth of the conditional is defined in relation to the closest antecedent world(s), not the closest consequent world(s). However, I see no reason to think that this asymmetry can capture the difference in standards of non-occurrence that we need.

We might think instead that the specification of an event may encode its non-occurrence conditions (as per option (ii)), but since one and the same event can be a cause in one relation, and an effect in another, it seems that the event itself cannot encode the two *different* standards that we require.

By far the most natural location for this variance is in the $\neg O$ clause (option (iii)) as this concerns the negation (non-) of the occurrence of the event. It is this option that I will pursue.

It is standard to treat P and $\neg P$ as exhaustive of the possibilities, and so Oc ("event c occurs") and $\neg O$ ("it is not the case event c occurs") should exhaust the possibilities too. This means that if event c can be taken to occur in a particular range of possible worlds, then $\neg O$ will be true in all and only the complement possible worlds. This fits with our CH standard: any world that lacks even a single essential feature must belong to the complement of the Oc worlds.

The $\neg O$ clause does not fit our EX standard, however. There is a modal gap between the worlds where event c occurs, and where it is completely excised: any world that has some of the

essential features of the event, but not others, will reside in this gap. Let us say that c *wholly* occurs in that set of worlds where all of c's essential features obtain (Oc), that c *partially* occurs in a strictly larger set of worlds (Pc) in which at least some of c's essential features obtain, and that c is *completely excised* in only those worlds where it does not even partially occur ($\neg Pc$). Thus, those worlds where c partially occurs, but does not wholly occur, i.e. the ($Pc \land \neg Oc$) worlds, constitute the modal gap between c's occurring, and its being completely excised.

This allows us to reflect the EX-CH standard of non-occurrence in a revised counterfactual conditional:

$$\neg Pc \,\square\!\!\rightarrow\, \neg Oe$$

To be read as: in all of the closest worlds where c does not even partially occur, e does not wholly occur.

Before moving on to consider the implications of this revision, it is worth noting the role that context is playing in this proposal. Context dictates what the essential features of the event are, and this in turn dictates what it is for the event to wholly occur or partially occur, and therefore what it is for the event not to occur, or not even partially occur. This leaves a story to be told about what features qualify to be *essential*[14], and how it is that context selects them. These are deep questions about the causal relata, and I won't attempt to deal with them here, but as far as I can tell my proposal is fairly neutral about exactly which treatment of events will fit. For example, everything I

---

[14] Lewis thought that events were predominantly intrinsic (1986, 264), and not overly disjunctive [p.266]. Nevertheless, he conceded that some minimal extrinsicality was required in specifying events—spatiotemporal location and the laws were both bound up in the identity condition for events [p.264].

have said here is consistent with taking events to be sets of regions of worlds (Lewis 1986), or as individuals within a counterpart theory.

On a different tack, some might complain that I have traced the non-occurrence issue to the wrong place and insist that context already dictates the ordering of worlds as part of the semantics of $\Box\!\!\rightarrow$. Thus when the cause event is to be completely excised, the similarity ordering of worlds will be such that the closest $\neg Oc$ worlds are those where it is completely excised. I think that this rival proposal needs to be outlined in more detail if it is to be properly assessed, but at first glance it does not seem to fit Lewis's (1979) similarity ranking, when considering the *causal* counterfactual. On Lewis's schema, we are to hold fixed the past as much as possible (no backtracking!), and deviate as little as possible from actuality (in terms of miracles and matters of fact) so as to render the antecedent true. But the antecedent in the original counterfactual conditional is $\neg Oc$, and the closest-to-actual way to realise that on Lewis's 1979 treatment, is to replace it with a near alternative, not completely excise it. So, either we propose a different antecedent, as I have done, or we deviate from Lewis's 1979 schema. As alluded to above, I think that there is good evidence that Lewis's schema has problems, but I see none to suggest that whatever replaces it will deliver the world ranking that would be required to settle the question of the non-occurrence of events.

Retrospective and Implications

I think it is illuminating to consider why this asymmetry in the standards of non-occurrence remained hidden for so long, and to ask: what are the implications of adopting it. I will consider each question in turn before summarising what the proposal in this paper amounts to. In the next section, I will turn to showing how a closely related issue affects other approaches to causation.

Retrospective

In retrospect it is easy to see how this ambiguity remained so well hidden. Firstly, the O representing 'occurs' is typically dropped in discussions of counterfactual analyses so the pivotal variable (as I have identified it) is almost always glossed over in the relevant literature. Second, English only has a limited range of natural locutions for specifying the key idea: the non-occurrence of the event, the event does/did not occur. These locutions do not track the distinction between the different standards of non-occurrence in question. Third, the two standards very frequently converge. This last point is worth demonstrating in more detail: when the cause event is specified in such a way as to evoke just a single essential aspect, it makes no difference whether you consider the non-occurrence of c to require that all of the essential aspects of c are absent (EX standard), or that just one is (CH standard). Since there is just one essential aspect the absence of one aspect or all amounts to the same thing and the two standards of non-occurrence are equivalent in that case. More generally, whenever the closest $\neg Oc$-worlds and $\neg Pc$-worlds are the same, the CH and EX standards of non-occurrence will give equivalent truth conditions for the causal counterfactual that takes c to be the putative cause.

Even when the closest $\neg Oc$-worlds and $\neg Pc$-worlds are different they may not bring about different truth values for the relevant counterfactual. For example, suppose that Suzy's throwing the ball causes the bottle to break and further suppose for illustration that the cause event in this context has just three essential features: it involves Suzy, a ball, and a throw. The complete excision of this event requires that there be no Suzy, no rock, *and* no throw. When c is completely excised, the window doesn't break and so the causal counterfactual comes out true on the EX standard of non-occurrence of the cause ($\neg Pc$). On the CH standard of non-occurrence for the cause, $\neg Oc$ fills the antecedent of the counterfactual. The worlds where c does not fully occur include those where Suzy

drops the ball, where someone else throws the rock and those where Suzy throws some other object. Of those $\neg Oc$-worlds, which is the closest? In this example it is plausible that the closest $\neg Oc$-worlds are those where Suzy drops the rock rather than worlds in which she morphs into a different person or where the rock is suddenly supplanted by some other object. If the closest $\neg Oc$-worlds (where Suzy drops) are all worlds where the window doesn't break, then the causal counterfactual will be true in that case. Of course I have manipulated a toy example here but the point to make is that there will be *some* examples where the cause will have a rich essence, and yet whether the EX and CH standards of non-occurrence are applied to the cause will make no difference to the truth of the causal conditional. This further explains why the distinction between the standards remained hidden.

Yet there remains a raft of cases where the complete excision policy for the cause is important. The three examples I introduced at the start concerning Socrates, Striker, and Michael Jackson for example, as well as the illustrative case of Suzy. In each case, replacing the putative cause with too near an alternative corrupted the intuitive results that the counterfactual analysis of causation seemed to deliver. Adopting an excision policy in general for the cause, *but not for the effect*, restores the intuitive verdicts.

Transitivity

One potential implication of the different standards of non-occurrence for the cause and for the effect, concerns transitivity. It was once standard to assume that causation was transitive, such that if c was a cause of d, and d was a cause of e, then c was a cause of e. In chains of causation, such as c—d—e above, the d event will be an effect in the first step, and a cause in the second. But if there is a different standard of non-occurrence applied to the cause and to the effect, then there is no guarantee that the feature of d that the complete excision of c alters, will have anything to do with

18

the difference the complete excision of *d* makes to *e*. In other words, transitivity should not be expected to hold in general.

Here is a schematic example. Suppose that event *d* has two essential properties (P and F), and that without *c*, *d* would have lacked just one of these (P). This is enough to render *c* a cause of *d* on the EX-CH standard. Now suppose that the other feature of *d* (F) makes a difference to *e*, such that the complete excision of *d* alters some essential feature of *e*. This is enough to render *d* a cause of *e* on the EX-CH standard. Thus, if causation were transitive, then *c* should be a cause of *e*, but nothing in this schematic case suggests that *c* has anything to do with *e*.

Of course, counterexamples to the transitivity of causation are nothing new—see McDermott (1995), Paul (2000), Hall (2000), Lewis (2004), Hitchcock (2001)—but the EX-CH amendment to the counterfactual analysis offers new insight into the structure of the issue, by being able to *predict* such examples purely on a schematic basis. See how the following counterexample (adapted from Ehring 1987, 323) fits with the schematic example above: throwing potassium salts into the fire (*c*) causes the flame to turn purple (*d*). The purple flame (*d*) causes the curtains to ignite (*e*). So, we have *c* causing *d* in virtue of making it purple (P) and *d* causing *e* in virtue of *d*'s being a flame (F), but we do not want to say that throwing the salts into the fire (*c*), caused the curtains to ignite (*e*). Thus, transitivity fails, just as we might expect once we notice the different standards that must apply to the cause and to the effect.

Proportionality

Counterexamples to transitivity that have the above structure appear to hinge on building too much irrelevant detail into the *d* event: surely the colour of the flame is irrelevant to the issue of whether the curtains catch fire! Such irrelevancy is permitted by the adoption of the complete excision

treatment of the cause, however, so if we are to rule out such irrelevancy in general when adopting a counterfactual approach to causation, we would need to abandon or amend the complete excision standard.

One such proposed amendment (though not typically applied to transitivity cases), is to add a proviso that causes should be *proportional* to their effects in the sense that no more tightly specified cause is *required* for the effect, and no less tightly specified cause is *sufficient* for the effect (paraphrasing from Yablo, 1992). Proportionality is supposed to ensure relevance. Here is a classic, and illuminating, example: Sophie the pigeon is trained to peck all and only red things. A scarlet patch is placed in front of Sophie and she pecks. Intuitively, it is better to cite the *redness* of the patch placed in front of Sophie as a cause of the peck, as specifying that it is *scarlet* may wrongly imply that a *crimson* patch wouldn't work, and merely saying that it is *coloured* may wrongly imply that a green patch would work too. Since citing the *scarletness* of the patch amounts to giving too much information, we should consider *the placing of the scarlet patch* to be a tighter specification of the event than is required, and so dis-proportionate to the effect. Similarly, since citing the *coloured* nature of the patch amounts to giving too little information, we should consider *the placing of the coloured patch* to be less tightly specified than is sufficient for the effect, and so also dis-proportionate. However, citing the *redness* of the patch gives just the right amount of information, just enough information, and not too much. Hence *the placing of the red patch* is the proportional cause of Sophie's peck.

Must *all* causes be proportional? List and Menzies (2009) think so: they adopt a view of causation on which only proportional causes are genuine causes.[15] Thus, on their view the placing of the red patch is a cause Sophie's peck (on certain contextual assumptions), but the placing of the

---

[15] List and Menzies actually take the proportionality principle, introduced by Yablo in (1992), to be *verified* by their independently reached difference making account, rather than a motivator for it. The directionality doesn't matter to my purposes here, however. See also Sartorio (2010).

scarlet patch is *not*. This is because on their approach the closest worlds in which the patch is not

scarlet, it is some other shade of red instead (List and Menzies 2009, 488).[16][17] Further, although they

do not discuss the case, their view entails that the purple flame does not cause the curtains to ignite,

but that the flame does. Again, this is because, on their approach, the closest worlds where the flame

is not purple, it is some other shade instead.

So, by adopting a proportionality constraint on causation, those counterexamples to

transitivity that rely on irrelevant detail in the middle place can be avoided. If proportionality is a

strong constraint on causation, in the sense that only proportional causes can be causes, then it stands

in direct conflict with the EX-CH approach outlined above. Does this speak against the EX-CH

account?

I say that it does not. The case in favour of adopting any counterfactual analysis of causation

is that it delivers intuition-matching results in ordinary cases. The EX-CH amended analysis does

well in this respect, as it agrees with common sense that the following are all true: Socrates' drinking

hemlock caused his death, Striker's low shot caused the keeper to dive, Suzy's throwing the heavy

red ball caused the bottle to break. The strong reading of the proportionality constraint, on the other

hand, undermines much of our ordinary causal reasoning. Since many other poisons would have

sufficed, citing the hemlock renders the claim about Socrates' death out of proportion. Similarly,

since a range of other shots would have made the keeper dive, specifying that it was low renders the

---

[16] It may prove important here that the List and Menzies approach is given in terms of *properties* rather than *events*. It seems much more natural to say "If the patch hadn't been scarlet, it would have been some other colour" than to say "if the placing of the scarlet patch had not occurred, a patch of a different colour would have been placed instead." Exploring this issue is beyond my scope here, however.

[17] Notice again that this approach of swapping in salient alternatives requires a CH reading of the antecedent in the counterfactual conditional. List and Menzies, and Sartorio both require such a reading.

whole causal claim out of proportion, and since any colour of heavy ball would have broken the window if thrown, citing the redness of it violates proportionality. If non-proportional causes are *not* causes, then each of these claims (and many more besides) are false.[18] A strong constraint on proportionality thus fails to accord with common sense.

I think the right response to this problem for proportionality is to step back from the claim that *all* causes are proportional, and instead see proportionality as some kind of optimal specification or as a *dimension of explanatory value* (to borrow Weslake's elegant phrase (2013)). But having given up on the strong interpretation, the proportionality constraint is no longer in conflict with the EX-CH approach: the first concerns the optimal form of a causal claim, the second concerns which claims are true. Further, on this preferred *weak* interpretation of the proportionality constraint does not rule out Socrates' drinking hemlock, the placing of the scarlet patch, or the purple flame, as causes, it just entails that there is some more optimal form of capturing those causal connections. This means that the weak interpretation of proportionality cannot resolve the problem cases of transitivity considered above (or do the work List and Menzies wish it to in respect of the causal exclusion argument[19]), as it cannot rule out as *false* any of the component claims.

Summing up

My discussion of the non-occurrence problem, and Lewis's *excision* proposal, is thoroughly conditionalised: *if* one were to adopt a Lewisian counterfactual analysis of causation, how should one deal with this issue? I have argued that Lewis's proposal requires refinement and revision to be at all

---

[18] A similar argument is given in (Bontly, 2005) and McDonnell (2017).

[19] I make the case for this in McDonnell (2017).

plausible, and that it leads, quite surprisingly, to a double-standard when it comes to the non-occurrence of causes and the non-occurrence of effects.

I have also argued that, if we are to be guided by our common sense assertions about ordinary causal cases (as is traditional), then we have to allow that some irrelevant detail in a causal claim does not render that claim false, contra strong interpretations of proportionality. The EX-CH approach neatly captures this aspect of our causal discourse, and helps us understand the workings of certain counterexamples to the transitivity of causation.

What has not been argued is that the EX-CH approach is a plausible theory of causation overall. There is nothing in this discussion that addresses long-standing issues about pre-emption cases, for example, or about the treatment of absences as causes. In fact, given that EX-CH undermines the transitivity of causation, it also undermines Lewis's treatment of early pre-emption cases which hinges on causation being transitive (Lewis, 1973). Those who reject such simple analyses of causation on such grounds, will find nothing to persuade them otherwise here.

Nevertheless, the ability to deal with the non-occurrence problem itself is no small matter. By way of a conclusion, I will turn to considering related problems for seemingly more capable variants from the difference making family of causal theories: contrastivism and interventionism.

Everyone's Problem

Contrastivism

Contrastive approaches to causation fall within the the same difference making family of causal theories as the original counterfactual analysis from Lewis. Both the contrastivist and the counterfactualist can agree that what makes one thing a cause of another is that the first *made a*

difference to the second, but they disagree on what the right test for difference making is. Where the counterfactualist takes an instance of causation to relate two events, the cause and the effect, the contrastivist takes it to relate four: the cause, the effect, the cause-foil, and the effect-foil.[20][21] The cause-foil is an alternative event to the cause, that is incompatible with the occurrence of the cause, and also *mutatis mutandis* for the effect. So, the cause-foil is *one* of the ways that the cause could have failed to occur.

Recall the extravagant examples I gave at the start to introduce the need for a *closeness* criterion when interpreting Lewis's counterfactual antecedent. Those examples demonstrated that there are certain ways for a cause *not to occur*, such that the effect still comes to pass (Socrates jumping from a tower, Striker throwing a grenade), and that there are certain ways for unrelated events to count as causes, such as when we suppose that instead of taking pain medication Jackson took part in an anti-Ronaldo conspiracy. These examples were intended to show what was *wrong* with some ways of filling in the antecedent in the counterfactual analysis.

The contrastivist approach has no such restriction, and its defenders embrace these possible antecedents, but deny that the causal verdicts they give are faulty.[22] It is true, the contrastivist will say, that Socrates' drinking hemlock caused his death, and it is true that it didn't. It did when the cause-foil is something like *drank no poison at all*, and it didn't when the cause foil is *jumped from tower*

---

[20] Not all contrastivists posit the quaternary relation, but I am persuaded by what Schaffer says (2005, 328) about why they should, so I stick with what I take to be the stronger version of the account.

[21] On this characterisation, I think Lewis's (2004) "Causation as Influence" account can be read as a proto-contrastivist theory. There, Lewis requires that we not just specify the events we are considering, but the range of alterations of those events that are salient too, when assessing a causal claim.

[22] See Northcott (2007, 121–22) for a clear elucidation.

instead. There is no contradiction here because the binary claim of the form '*c* caused *e*' is semantically incomplete, and once the causal claim is completed by including the cause-foil and the effect-foil, there is no contradiction. Certainly, one claim might be more relevant than the other in a particular context, but that is not to say that the other is thereby false.

It is thus straightforward to create "counterexamples" to contrastivism: take any paradigmatic non-cause and pair it with a foil which would have impacted the effect (nuclear holocaust is an easy plug-in, for example). The contrastivist will consider the non-cause a cause. Or, take a paradigmatic cause and pair it with a foil which would have had exactly the same impact on the effect (a super-close alternative, for example). The constrastivist must deny that the cause is a cause. However, these are not counterexamples as such, since the paradigm claims in each case—that *c* was not a cause in the first, that *c* was a cause in the second—are semantically incomplete until they are paired with a foil, and whatever foil we take in the paradigmatic case will be different from the foil that drives the counterexample. There will be no conflict. This will be a quite general strategy for the contrastivist to rebut seemingly unpalatable binary claims that they might be committed to.

The contrastivist is thus an *invariantist* contextualist about causal claims: context will often be required to help complete the causal claim as uttered, but the causal truths—once semantically disambiguated—will not vary with context at all. What does vary with context is the relevance of any given causal truth to the conversation, and so it can be relevant that Socrates's drinking hemlock rather than wine caused his death, but it is not relevant that his drinking hemlock rather than jumping from a tower did not. Notice, though, that this simply moves the question of what the relevant alternative to the cause or effect is from the semantics, where it is the non-occurrence problem I outlined for the counterfactualist, to the pragmatics, where it becomes a question about

25

which foil is relevant in the context. The contrastivist therefore faces the non-occurrence problem in a different form: not "what is it for an event not to occur?" but rather "what is the relevant way in which the event would fail occur?"[23] Schaffer discusses this question in his (2012)(esp. p.55-57), and concludes that it is far from obvious how a single context can provide the *two* foils that the contrastivist requires.

Interventionism

Interventionism also falls within the difference making family of causal theories. The canonical expression, from Woodward (2003), defines several causal notions in terms of whether or not an idealised *intervention* upon one variable, would result in changes in another. Thus interventionism is also a variety of counterfactual theory.

The differences from Lewis's account, however, are many. For the purposes of this paper, the following are most important: interventionism takes the relata of causation to be *variables* broadly conceived, not just events, and interventionists reject Lewis's similarity semantics for counterfactuals. The result of these divergences is that interventionists appear to escape the non-occurrence of events problem as outlined, since that problem first concerns events, and second arises from issues concerning the modal 'closeness' of alternates. I will argue here that they have not genuinely avoided the non-occurrence problem, but rather relocated it.

---

[23] In his (2004), Lewis proposes that we define causation in terms of influence, and influence in terms of counterfactual dependence between "substantial range[s]" of "not-too-distant" (p. 91) alterations of the cause and effect events. Given the reliance on sets of alterations, and the role of salience in determining which of the pairings are relevant to a particular causal enquiry, I think it is fair to bundle this later-Lewis with the contrastivist: the non-occurrence issue has simply been moved to the pragmatics.

When the Lewisian counterfactualist considers what Socrates might have done, had he not drank hemlock, they look to the closest world to actuality in which Socrates did not drink hemlock. The bulk of this paper has been devoted to the issues that arise from that. The interventionist, on the other hand, looks to their causal model see what other values of the relevant variable there might be. For example, if the variable is *Socrates's Drinking* then the values of that variable might include *drinking hemlock*, *drinking arsenic*, *drinking wine*, *drinking nothing*. For the interventionist, Socrates's drinking hemlock caused his death if there is some other value that the relevant variable could have taken such that Socrates did not thereafter die. So, since drinking wine, and drinking nothing, are possible values the variable could have taken instead, and supposing that these would not have resulted in Socrates death shortly thereafter, then it is true by interventionist lights to say that Socrates's drinking hemlock caused his death.

It should be clear enough why *drinking coke* or *drinking battery acid* aren't values that the interventionist might include in their model—they are not relevant alternatives for Socrates—but what is not clear is what makes any particular set of variables relevant, or what alternate values of those variables should be included. This matters. Suppose that the only values in the model were *drinking hemlock* and *drinking arsenic*. In that model, it is false to say that had Socrates not drunk the hemlock, he wouldn't have died, and so *relative to that model* the canonical causal claim about the cause of Socrates's death is false.

This is what we might call the *variables problem* for interventionism, following recent work by Franklin-Hall (2016). The problem, in a nutshell, is that we must first work out what variables, and values of those variables, we are interested in, before we can build a model which captures the causal facts. The non-occurrence problem concerned the relevant alternatives for events in a simple

27

counterfactual theory, and the variables problem concerns the relevant alternatives for values of variables within a causal model. The core problem of relevant alternatives, is common.

So, the counterfactualist locates the *relevant alternative* problem in the semantics of the counterfactual, the contrastivist locates it in the pragmatics of causal assertions, and the interventionist locates it in the construction of the model. I contend that this is a problem that has thus been moved, but not solved.

Acknowledgements

References

Bennett, Jonathan. 2003. *A Philosophical Guide to Conditionals*. Oxford University Press.

Bontly, Thomas D. 2005. "Proportionality, Causation, and Exclusion." *Philosophia* 32 (1-4). Springer Netherlands: 331-48.

Ehring, Douglas. 1987. "Causal Relata." *Synthese* 73 (2): 319-28.

Franklin-Hall, L. R. 2016. "High-Level Explanation and the Interventionist's 'Variables Problem'."

The British Journal for the Philosophy of Science 67 (2): 553-77. doi:10.1093/bjps/axu040.

Hall, Ned. 2000. "Causation and the Price of Transitivity." *Journal of Philosophy* 97 (4): 198-222.

Hawthorne, John. 2005. "Chance and Counterfactuals." *Philosophy and Phenomenological Research* 70 (2).

Wiley-Blackwell: 396-405.

Hájek, Alan. manuscript. "Begging to Differ with Similarity Accounts of Counterfactuals."

Hitchcock, Christopher. 2001. "The Intransitivity of Causation Revealed in Equations and Graphs."

The Journal of Philosophy 98 (6). Journal of Philosophy, Inc.: 273-99.

Kim, Jaegwon. 1973. "Causation, Nomic Subsumption, and the Concept of Event." *Journal of*

Philosophy 70 (8): 217-36.

Kment, Boris. 2006. "Counterfactuals and Explanation." *Mind* 115 (458). Oxford University Press:

261-310.

Krasner, Daniel, and Mark Heller. 1994. "The Miracle of Counterfactuals: Counterexamples to

Lewis's World Ordering." *Philosophical Studies* 76 (1). Springer: 27-43.

Lewis, David. 1973. "Causation." *Journal of Philosophy* 70 (17): 556-67.

———. 1979. "Counterfactual Dependence and Time's Arrow." *Noûs* 13 (4): 455-76.

———. 1986. "Events." In *Philosophical Papers*, II:241-69. Oxford University Press.

———. 2004. "Causation as Influence." In *Causation and Counterfactuals*, edited by John Collins, Ned

Hall, and L. A. Paul, 75-106. The MIT Press.

List, Christian, and Peter Menzies. 2009. "Nonreductive Physicalism and the Limits of the Exclusion Principle." *Journal of Philosophy* 106 (9). 475-502.

McDermott, Michael. 1995. "Redundant Causation." *British Journal for the Philosophy of Science* 46 (4): 523-44.

McDonnell, Neil. 2016. "Events and their counterparts." *Philosophical Studies* 173: 1291.

———. 2017. "Causal exclusion and the limits of proportionality." *Philosophical Studies* 174: 1459.

Northcott, Robert. 2007. "Causation and Contrast Classes." *Philosophical Studies* 139 (1): 111-23.

Paul, L. A. 2000. "Aspect Causation." *Journal of Philosophy* 97 (4): 235-56.

Paul, L. A., and Ned Hall. 2013. *Causation: A User's Guide.* Oxford University Press.

Sartorio, C. (2010). "The Prince of Wales Problem for Counterfactual Theories of Causation." In A. Hazlett (Ed.) New Waves in Metaphysics, (pp. 259-276). Palgrave McMillan, New York.

Schaffer, Jonathan. 2004. "Counterfactuals, Causal Independence and Conceptual Circularity." *Analysis* 64 (4): 299-308.

———. 2005. "Contrastive Causation." *Philosophical Review* 114 (3): 327-58.

———. 2012. "Causal Contextualisms." In *Contrastivism in Philosophy: New Perspectives*, edited by Martijn Blaauw. Routledge.

Weslake, Brad. 2013. "Proportionality, Contrast and Explanation." *Australasian Journal of Philosophy* 91 (4): 785-97.

Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press.

Yablo, Stephen. 1992. "Mental Causation." *Philosophical Review* 101 (2): 245-80.