
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/149495/

Deposited on: 11 October 2017

Enlighten – Research publications by members of the University of Glasgow
http://eprints.gla.ac.uk
MAKING A CONTRIBUTION AND MAKING A DIFFERENCE

Neil McDonnell

ABSTRACT

There are at least two different concepts that philosophers might target when analysing causation: a pre-selective notion, and a selective notion. This paper argues that these two distinct conceptions have been conflated to date, citing the puzzles of overdetermination, extensionality, and transitivity as evidence. The primary aim of the paper is to help re-set the methodological scene concerning analyses of causation.

There are at least two target concepts philosophers could be aiming for when analysing causation: a pre-selective notion that ignores our pragmatic focus on what is relevant or salient, or a selective notion which takes this focus, clearly present in our ordinary causal discourse, into account. In this paper I will motivate and defend the thesis that these two targets have been widely conflated in the causation literature.

I will first argue that Lewis’s influential rejection of the so-called fragility strategy for dealing with late pre-emption cases precipitated a shift in methodology and, by extension, a shift in the target concept under consideration. I think that this equivocation is widespread, and I will present three puzzles from the causation literature—concerning overdetermination, extensionality, and transitivity—as evidence. In each case, positing an equivocation in the target concept, illuminates the puzzle. I will conclude from this that two concepts are at play in this literature, and that they need to be carefully distinguished.

This point stands regardless of what one thinks of the philosophical value of the target concepts, or how they relate to one another. I will briefly discuss the options at the end, but in this paper I will remain entirely neutral on which is the right target. The argument given in this paper is intended to re-set the methodological scene.

METHODOLOGY AND LATE PRE-EMPTION

In his original presentation of the counterfactual analysis of causation, Lewis characterised his target as follows:

We sometimes single out one among all the causes of some event and call it “the” cause, as if there were no others. Or we single out a few as the “causes,” calling the rest mere “causal factors” or “causal conditions.” Or we speak of the “decisive” or “real” or “principal” causes. We may select the abnormal or extraordinary causes, or
those under human control, or those we deem good or bad, or just those we want to talk about. I have nothing to say about these principles of invidious discrimination. I am concerned with the prior question of what it is to be one of the causes (unselectively speaking). My analysis is meant to capture a broad and non-discriminatory concept of causation. (Lewis 1973, 558–59)

This statement is doubly telling in respect of Lewis’s methodology. It says in the first instance that we will not be able to analyse the target concept by reading off of our ordinary linguistic practices, given that these practices are shot through with the sort of pragmatic selectivity that Lewis wishes to look beyond. Secondly, though, it licenses our analysis to count more causes than we might ordinarily think there are, but not less. If we countenance a cause in our ordinary talk, then Lewis’s analysis had better be able to count it too. The reverse does not hold.

Hall later echoes this methodological stance, under a different heading. In his own words, he is aiming for an egalitarian concept of causation that does not privilege foreground over background conditions on the basis of salience (2004, 228). Furthermore he warns against the cheap rejection of counterfactual theories by ignoring this egalitarian requirement:

Suppose that my favorite analysis counts the big bang as among the causes of today’s snowfall (a likely result, given transitivity). How easy is it to refute me, by observing that if asked what caused the snowfall [...] we would never cite the big bang! Of course, the right response to this “refutation” is obvious: It conflates the egalitarian sense of “cause” with a much more restrictive sense (no doubt greatly infected with pragmatics) that places a heavy weight on salience. (Hall 2004, 228)

With just such a pre-selective target in mind, Lewis offered the following counterfactual analysis of causation:

\[c \text{ causally depends on } e \text{ iff } \neg Oc \rightarrow \neg Oe.\]

In English: \(c\) causally depends on \(e\) if and only if it is the case that \(had c\) not occurred, \(e\) would not have occurred, and \(c\) is a cause of \(e\) if and only if there is a chain of such dependence between them.

This analysis works in a wide range of ordinary cases, and I suspect that it is for this reason that this simple counterfactual analysis remains of interest long after even Lewis himself abandoned it. I turn now to a class of cases that he thought unanswerable on his original theory, the cases of so-called late pre-emption.

**LATE PRE-EMPTION**

Here is a canonical example of late pre-emption:

**LP:** Billy and Suzy are out to vandalise and they each pick up a rock, and throw it accurately at a window at the same time. Suzy throws harder, though, and her rock reaches the window first. The rock strikes the window, the window breaks, and then Billy’s rock passes harmlessly through the void a moment later.

In this case it is clear that Suzy caused the window to break, but we are to suppose that Billy’s rock had sufficient heft to break the window, had it not been broken already. As such the crucial counterfactual
for Lewis is false. It is false because absent Suzy’s throw, Billy’s rock will break the window anyway: if Suzy hadn’t thrown the rock, the window would not have broken. The guarantor undermines the dependence.

At first glance, pre-emption problems appear to admit of an obvious solution: since the window would have broken slightly later as a result of Billy’s throw, the window breaking that would have occurred (call it $e_2$) is a different event from the one that did (call it $e_1$). So, the one that did occur ($e_1$) depended on Suzy since without her that event would not have occurred.

However, Lewis, and many others since, have rejected this solution. The problem is, they say, that taking events to be so sensitive in terms of exactly when and how they occur—to treat events as modally fragile, to use the jargon—opens the causal floodgates. Imagine that a dog is barking in the distance as Billy and Suzy throw their rocks. We know from our present-day physics that this dog’s bark will create an imperceptible vibration in the window, altering the precise nature of its breaking. If our individuation criteria for events are so precise that any minor variation in the timing and manner of the event makes it a new event, then the window breaking event will counterfactually depend on this dog’s bark. Combined with a counterfactual theory of causation, this renders the dog’s barking a cause of the window breaking event. Lewis considered this a reductio of the fragile events strategy (1986b, 198).

I think that this argument marks a methodological shift. By hypothesis, the dog’s barking does indeed make some difference to the window breaking event since without the barking, the breaking would have been in some way different. The problem with citing the dog’s bark, is that it does not make the difference that we are interested in at the human level, or one that we are morally or practically invested in. It just isn’t relevant. But the methodology Lewis and Hall outlined warned us not to be swayed by issues of relevance: we are supposed to be identifying all of the causes, not just those that are salient.

There is a rival interpretation, however, and one that may seem more charitable when applied specifically to Lewis. Suppose that we take events to be transworld sets of regions (containing no more than one region per world), as per (Lewis 1986a). On this view the window breaking, when modally fragile, is quite literally a different event from the window breaking when even slightly more modally robust. On this extremely fine-grained conception of events, there is a window breaking event which the dog-bark causes, and there is a window breaking which it does not, and so Lewis can insist that the latter event is not caused by the dog’s bark even on the most egalitarian conception of ‘cause’.

I think this probably is a more charitable interpretation in terms of what Lewis wrote about events, but I do not think that it is more charitable overall. First of all, such fine-graining of events would require that we count vastly more events in any given region than common sense could make sense of: one for every different such transworld set of which the actual region was a member. Second, in attempting to figure out which of this multitude of events was being picked out in a given causal claim, Lewis’s own interpretative metric (1983) would require that we alight on the interpretation that rendered the claim true. If there genuinely are so many events to choose from when considering the referent of ‘window breaking’, then there will precious few causal claims that are false, since there will always be some transworld set which would render them true. We certainly should not think that the claim about the
dog’s bark being a cause of the window’s breaking are to be dismissed tout court if there exists some event which would make it true.

Whilst I think it may be more charitable to diagnose a methodological shift in Lewis than to attribute such a view, my task here is not Lewis exegesis, and so it is not the place to try and settle such disputes about his particular treatment of events (though, see Kaiserman 2016, McDonnell 2016, and Vance 2016 for recent, and relevant, critique). Whilst few have adopted such a super-fine treatment of events (Paul 2000, and Schaffer 2005, both adopt a significantly more coarse grained treatment, for example), many have followed Lewis’s reasoning that we should reject the fragility strategy because it counts trace influences as causes. So, even if Lewis isn’t guilty of a methodological shift in virtue of individuating events as he does, others who do not adopt that view of events will be. Thus, I will restrict my diagnosis of a methodological shift to those philosophers who, (i) seek a pre-selective or egalitarian analysis, and (ii) (I think sensibly) do not retreat to a super-fine conception of events, but (iii) still complain that the dog’s bark should not count amongst the fullest set of causes.

Let us say, then, that there are two target concepts that we might aim for. The first is the genuinely pre-selective, egalitarian and wholly objective notion of making some difference, such that any contributing factor will do. For ease, I’ll refer to this as making a contribution or MAC. The second is the interest-relative notion of making some particular difference, such as making the difference between the window breaking and not breaking, or between the window breaking as it did and breaking very slightly differently. The particular difference in question may be weakly objective in the sense that once we fix the relevant or salient difference in a given context, it is then an objective matter which prior events made that difference. But in a stronger sense it is not objective, as it renders causation relative to contexts or interests. I’ll refer to this notion as making a difference or MAD on the understanding that it means to make a particular difference, the salient difference in some context, rather than just any difference at all.

In the discussion of late pre-emption cases it seems to me that the target concept has shifted from that of making a contribution to that of making a difference, and in the process set the tone for subsequent debates. It is my contention in this paper that the shift led philosophers to conflate the two target notions when considering a range of important and challenging issues.

If this contention is correct, then we should expect to see some evidence of conflict in cases where the MAC and MAD conceptions yield different verdicts. It is to such cases that I now turn.

**Making a Contribution and Making a Difference**

In this section I will consider three issues that have arisen in the literature on counterfactual accounts of causation, each of which presents us with something of a puzzle. In each case, I aim to show that distinguishing the MAC and MAD conceptions of causation helps us understand the puzzle, though not necessarily resolve it.

**Overdetermination**
The cases of late pre-emption discussed above fall within the broader set of cases of so-called redundant causation. Pre-emption cases are asymmetric versions of redundant causation, because there is a clear asymmetry between the cause (Suzy) and the non-cause (Billy) to which the counterfactual analysis is blind. Symmetrical cases, by comparison, seem to have two equal candidates to be a cause and, as with pre-emption, the presence of two sufficient causes for some effect is taken to undermine the counterfactual analysis of causation, which delivers the result that neither is a cause. A familiar example will help illustrate:

**OD:** A firing squad of eight shoots and kills a prisoner. Each squad member’s bullet would have been deadly on its own.

In such a case, it is obvious that the firing squad killed the prisoner, but on a counterfactual account of causation none of the members of the squad is a cause because the death did not depend on them individually given the other lethal shots. According to the simple counterfactual analysis, though, the firing squad as a whole did kill the prisoner, since had the firing squad (in its entirety) not shot, then the prisoner would not have died.

Is the verdict conferred by the counterfactual analysis the correct one? Lewis points out that we only have intuition to guide us, and that his intuition does not confer a clear judgement about such cases. He concluded that the correct result in overdetermination cases should be ‘spoils to the victor’ (1986b, 194): whatever verdict the best all-else-considered theory of causation determines, should be considered the correct one.

This response to overdetermination cases has left many unsatisfied, but I will not attempt to adjudicate the matter here. Instead I will show that the puzzling nature of these cases can be explained if there are two alternative concepts of causation being conflated.

On the MAC conception, it is clear that every shooter in the squad made some contribution to the death. There were after all eight bullet holes in the body, not to mention the reverberation from eight guns firing and the force imparted to the body by eight bullets striking. It is clear that the MAC conception of causation considers every shooter a cause. A closely related line of thought has been pursued by (Bunzl 1979), who argued that given the physics of our world, genuine overdetermination is impossible here and so is not a pressing philosophical problem. I think it is clear that Bunzl was operating on what I have identified as the MAC conception of causation.

On the MAD conception, by comparison, it is clear that no one member of the squad made the difference between the prisoner dying and not dying, since we are told in the set up that each shot was sufficient for the death. Of course, nothing could make the difference between a mortal dying and not dying, since the outcome is always guaranteed eventually. I think this just goes to show that, when Lewis’s theory says that the firing squad did, but that the individuals did not, cause the death, there is at least some implicit limitation being applied on the range of deaths that are being considered salient, and those which are not. In other words, the problem of overdetermination, such as it is, occurs at the level of salience: there is a salient difference that the firing squad made, but that the individual members of it did not.
It would make sense of the puzzling intuitions of this case, then, if the MAC and MAD conceptions were being conflated in general. On the one conception, every member of the squad is a cause, and on the other, and with salience playing a crucial role, they are not. I think we can see this as progress in understanding the issue of overdetermination, without having to declare that either conception is correct.

CONTEXTUALISM AND EXTENSIONALITY

Menzies (2009) identifies what he takes to be a widely held idea: that causation is a natural relation between events. Here ‘natural’ is not supposed to call to mind some specific metaphysical thesis, but rather is supposed to capture the idea that causation is in some sense mind- and interest-independent, and out there in the world. A consequence of thinking of causation as just such a natural relation, is that it should not matter how we choose to describe the events in question, the same causal truths ought to hold. As Menzies points out, if causation is a natural relation in this sense, then:

...causal propositions should be extensional in the sense that the substitution of one event nominal for a second coreferential event nominal in a causal proposition should preserve the truth-value of the proposition. (Menzies 2009, 346)

The idea that the relation of causation is extensional has substantial pedigree but seems at odds with the growing consensus that the truth of causal attributions can vary when co-refering terms are substituted in. Here I recount a well-worn example:

1. McEnroe’s tension caused him to serve awkwardly.

2. McEnroe’s tension caused him to serve.

If causal propositions are extensional, and on the assumption that we are referring to a single serve, then whether or not it is referred to as awkward, the truth value of the causal proposition should not change. However, we would readily accept the truth of 1 and yet deny 2. Intuitively, McEnroe’s tension made a difference to the type of serve he engaged in, not whether or not he served at all.

Once we take such cases seriously, it looks like we have a conflict between the idea of causation as an extensional relation, and the fact that our causal talk appears to generate an intensional context. If Menzies is right that conceiving of causation as a natural relation entails that causal propositions should be insensitive to nominal substitution, then our causal talk is in conflict with the idea that causation is just such a natural relation. This is a puzzle.

I think the puzzle is explained, however, if we consider the issues in light of the distinction between making a contribution and making a difference. It should be clear that McEnroe’s tension makes some contribution to the timing and manner of his serve. So, by the lights of the making a contribution standard for causal connectedness, both 1 and 2 are true. In fact, the MAC conception will consider the tension to be a cause no matter how the effect is described: his third serve, his worst serve, etc. Yet his tension did make one specified difference to the serve—it made it awkward rather than graceful—without making another—the difference between serving and not serving. On the hypothesis that the
different descriptions in 1 and 2 imply these alternatives respectively, then it is clear that on the MAD conception there is a truth value shift when the event nominal is switched.

In this case, the MAC conception gives results that are consistent with the extensionality of causation, but the difference making conception does not. On the other hand, the difference making conception tracks the shift in the intuitive truth values from 1 to 2, whereas the making a contribution conception cannot account for that shift.

We can conclude from this example that binary causal propositions are not extensional, if interpreted in line with the MAD conception. We can also conclude that the MAC conception does not perfectly track our ordinary causal ascriptions. This is evidence for the hypothesis that neither conception alone can account neatly for both the causal platitude that causation is a natural relation, and our intuitions in ordinary cases. The puzzle comes from trying to explain both within a single conception of causation.

**Transitivity**

Is causation transitive? If $c$ is a cause of $d$, and $d$ is a cause of $e$, does that entail that $c$ is a cause of $e$? Lewis certainly thought so, and built transitivity into his original account. Hall is similarly explicit. Transitivity is, he says, a bedrock datum, a central connotation of our causal concept. However, take the following example:

While skiing, Suzy falls and breaks her right wrist. The next day, she writes a philosophy paper. Her right wrist is broken, so she writes the paper with her left hand[...] She writes the paper, sends it off to a journal, and it is subsequently published. Is Suzy’s accident a cause of the publication of the paper?

Of course not. (Paul 2000, 205)

Here the problem is that whilst we have no issue with saying that the accident ($c$) caused the left-handed writing ($d$), and that writing caused the publication ($e$), we find that the result entailed by transitivity is absurd: the accident ($c$) did not cause the publication of the paper ($e$). Given the antecedent plausibility of the transitivity of causation, such counterexamples present another puzzle, and once again I think that the two conceptions of causation that I have been discussing can help shed some light.

On the assumption that the writing and the left-handed writing were one and the same event ($d$), then it is clear that the accident made some difference to this event, and it is evident that the publication would not have occurred if the writing hadn’t. It should be clear then that $c$ made a contribution to $d$ and that $d$ made a contribution to $e$. However, as Paul points out, the difference made in the first causal step from $c$ to $d$ is a difference in which hand did the writing (the left-handed aspect of the event). But in the second step from $d$ to $e$, the publication is altered not by the left-or-right handedness of the writing (which $c$ did influence), but rather by whether there was any writing at all (which we suppose that $c$ did not). On the MAD conception of being a cause, the accident makes a particular difference to the writing, and a different feature of the writing—its quality, presumably—causes the subsequent publication. When we specify which particular difference we are interested in at each step, we can see that there is a mis-match in the difference in $d$ which is salient when it is the effect, and the difference
which is salient when it is the cause. Paul proposes that we only commit to the transitivity of causation when the relevant aspect of $d$ is held fixed across each step. In terms of the MAD conception, this would amount to saying that MAD is transitive only when the difference in $d$ is the same in each step. If that is the case, then the above scenario is no counterexample to the transitivity of MAD, since whilst the conclusion remains unpalatable, the transitivity of MAD does not license that conclusion.

I contend that it is the MAC conception, insensitive to which difference was made, so long as some difference is, that makes such ‘counterexamples’ initially compelling. That is no surprise, given that MAC will not be transitive in general: $c$ could make a difference to the colour of $d$, but it could be $d$’s momentum that alters $e$. Absent some other mechanism, $c$ will make no contribution to $e$ and so MAC is not transitive.\footnote{11} The transitivity of the MAD conception, on the other hand, does not seem to be impacted by these examples once we take on board Paul’s lesson about equivocating on the middle place of the chain. Thus, in these cases, positing the conflation of the MAC and MAD conceptions helps explain the puzzle.\footnote{12}

**CONCLUSION**

The foregoing discussion aims to show that there are two importantly different target concepts that we might aim to analyse in the philosophy of causation. I have defended the hypothesis that these concepts are being conflated in discussions around late pre-emption, overdetermination, extensionality and transitivity, and that distinguishing them helps us understand those puzzles better.

What I have not done, is make any claim about is the philosophical value of either concept, or their relation to one another. If we were to set the notions against one another as rivals for the title ‘causation’, for example, then many would doubtless complain that the MAC conception is far too liberal a notion to fit the various roles that causation is taken to play in ordinary usage.\footnote{13} On the other hand, Strevens (2008, 181) has argued that our ordinary causal assertions of the form ‘$c$ caused $e$’ should be read as giving causal explanations, and so the (lack-of) assertability of such claims should not be taken to strictly limit our analysis of causation itself. Further, Strevens defends a view of causation that is extensionally equivalent to the MAC conception, and suggests that this fits with the permissive concept of causal relevance that (Salmon 1997) relied upon in giving his theory of causal explanation (Strevens 2008, 42). A more conciliatory approach would be to synthesise both conceptions into a single theory of causation, in the manner Hall and Paul allude to in the conclusion of their recent book:

Although we have focused mostly on the ontological part of the project, we like the idea of a mixed approach for a final theory. Such an approach would be sensitive to the need to give a suitably objective account of the causal structure of the world, yet recognise that any account that can do justice to our causal intuitions will have to have some pragmatic features drawn from our ordinary causal judgements, at least enough of some features to accommodate our causal explanations. (Paul and Hall 2013, 254)

One need not take a particular stand on how we ought to analyse or interpret the presence of these two notions in the philosophical literature, in order to appreciate that there are two notions at work. Howsoever one chooses to proceed, this distinction has methodological consequences: those
attempting to analyse the MAD conception may reasonably explain away certain intuitions or platitudes as vestiges of the MAC conception\(^4\); those wishing to analyse the MAC conception need not accept our ordinary causal assertions as the litmus test for a successful account. More generally, philosophers of causation should take care to match the target under consideration with the appropriate methodology for investigating it.

University of Glasgow

---

NOTES

My thanks to audiences from the Joint meeting of the North and South Carolina Philosophy Societies in Boone (2016), workshop on Language and Ontology at the Humboldt University of Berlin (2016), and the Emergence, Exclusion and Causation workshop at the University of Glasgow, 2016, for helpful feedback. Thanks also to Umut Baysan, Michael Clarke, Stephan Leuenberger, Nathan Wildman, and Richard Woodward for comments and suggestions. Particular gratitude is due to Laurie Paul who, philosophical disagreements notwithstanding, saw a paper here before I did.

---

1 Most explicitly in (Lewis 2004).

2 For brevity I ignore early and trumping varieties here.

3 Lewis would of course claim that the fragile and robust alternates are not distinct in his sense, but they are nevertheless non-identical. For details see (Lewis 1986a, 256).


5 It is worth noting that the fragility strategy may not be a general solution to pre-emption problems—a point trumping cases are intended to demonstrate (Schaffer 2000)—but whether it is or not is orthogonal. My point here is a purely about a shift in methodology.

6 What counts as a resolution will depend in part on what we take the success criterion to be, and that is in part what is at question in this paper. I therefore stick to the weaker claim to avoid unnecessary controversy.


8 It is important to note that Menzies was identifying this ‘platitude’, not endorsing it.

9 See, inter alia, (Anscombe 1993; Davidson 1967; Levin 1976; Sher 1974).

10 Some will insist that there are two events here, not one. I disagree (see McDonnell 2016), but I need not argue the case here. Examples where the two statements only differ in stress are problematic for
such fine-grainers too: Socrates’s drinking hemlock at dusk caused his death versus Socrates’s drinking hemlock at dusk caused his death. For fuller discussion see Achinstein (1975), Schaffer (2005, 307–9), and Menzies (2009, 348).

11 It is a quirk of our relativistic physics that no true counterexample should be expected in our world. As Bunzl pointed out: everything in the backwards light-cone of e makes some contribution to it via gravity. Thus for any actual chain of contribution from c to d to e, c will make some contribution to e.

12 There are further cases to consider for the defender of transitivity—cases of switching, and short circuits in particular. This does not affect my point that there are two concepts at play in the mis-match cases. I am not suggesting that the distinction alone gives us a general solution to the problem of transitivity.

13 Schaffer alludes to such a response in his (2014), for example.

14 I read Woodward (2003) as taking this approach. The pragmatic factors that influence the choice of variables and values for causal models place his Interventionist approach clearly in the MAD camp. Furthermore, Woodward is unapologetic about rejecting certain standard metaphysical assumptions: that causation can be reductively analysed, or that causation is transitive.

REFERENCES


