Causation as Influence
Author(s): David Lewis
Source: *The Journal of Philosophy*, Vol. 97, No. 4, Special Issue: Causation (Apr., 2000), pp. 182-197
Published by: Journal of Philosophy, Inc.
Stable URL: http://www.jstor.org/stable/2678389
Accessed: 01/09/2008 15:06

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=jphil.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

http://www.jstor.org

CAUSATION AS INFLUENCE*

I have long advocated a counterfactual analysis of causation. But the simplest counterfactual analysis breaks down in cases of redundant causation, wherefore we need extra bells and whistles. I have changed my mind once more about how those bells and whistles ought to work.[1]

I. PREEMPTION REVISITED

It sometimes happens that two separate potential causes for a certain effect are both present; and either one by itself would have been followed by the effect; and so the effect depends upon neither. Call any such situation a case of *redundant causation.* Some cases of redundancy are symmetrical: both candidates have an equal claim to be called causes of the effect. Nothing, either obvious or hidden, breaks the tie between them. It may be unclear whether to say that each is a cause or whether to say that neither is a cause (in which case we can still say that the combination of the two is a cause). But, anyway, it is out of the question to say that one is a cause and the other is not. Because it is unclear what we want to say, these symmetrical cases are not effective test cases for proposed analyses of causation. Set them aside.

Other cases are asymmetrical. It is very clear what we want to say: one of the two potential causes did cause the effect, the other one did not. Call the one that did a *preempting* cause of the effect; call the other one a *preempted* alternative, or backup.

When our opinions are clear, it is incumbent on an analysis of causation to get them right. This turns out to be a severe test. The simplest sort of deductive-nomological analysis flunks: the preempted alternative is a member of a minimal set of conditions lawfully sufficient for the effect, yet it is not a cause. The simplest sort of counterfactual analysis likewise flunks: the preempting cause is not a condition without which the effect would have been absent, yet it is a cause. Both these attempts fail because they treat the preempting

* This article is an abridgement of the Whitehead Lectures, Harvard University, March 1999. Thanks are due especially to the other four authors whose articles appear in this issue of this JOURNAL; I have learned so much from them that this could almost have been presented as a jointly authored paper, were it not that the four of us continue to disagree extensively. Thanks are due also to Jonathan Bennett, Christopher Hitchcock, Simon Keller, Stephanie Lewis, Cei Maslen, D.H. Mellor, D.H. Rice, and the Boyce Gibson Memorial Library.

[1] Here I shall confine my attention to causation under deterministic laws. More likely than not, causation in the actual world requires a probabilistic analysis, but that raises problems independent of those I shall be discussing.

0022-362X/00/9704/182–97 © 2000 The Journal of Philosophy, Inc.

cause and its preempted alternative alike, whereas we know very well that one is a cause and the other is not. A correct analysis will need to discern the source of the difference.

II. TRUMPING

I used to think that all cases of preemption were cases of *cutting*: cases in which, first, there is a completed causal chain running from the preempting cause all the way to the effect; but, second, something cuts short the potential alternative causal chain that would, in the absence of the preempting cause, have run from the preempted alternative to the effect. Some think so still, but I have learned better.[2]

The sergeant and the major are shouting orders at the soldiers. The soldiers know that in case of conflict, they must obey the superior officer. But as it happens, there is no conflict. Sergeant and major simultaneously shout 'Advance!'; the soldiers hear them both; the soldiers advance. Their advancing is redundantly caused: if the sergeant had shouted 'Advance!' and the major had been silent, or if the major had shouted 'Advance!' and the sergeant had been silent, the soldiers would still have advanced. But the redundancy is asymmetrical: since the soldiers obey the superior officer, they advance because the major orders them to, not because the sergeant does. The major preempts the sergeant in causing them to advance. The major's order *trumps* the sergeant's.[3]

We can speculate that this might be a case of cutting. Maybe when a soldier hears the major giving orders, this places a block somewhere in his brain, so that the signal coming from the sergeant gets stopped before it gets as far as it would have if the major had been silent and the sergeant had been obeyed. Maybe so. Or maybe not. We do not know one way or the other. It is epistemically possible, and hence it is possible simpliciter, that this is a case of preemption without cutting.

III. COMMONPLACE PREEMPTION

Trumping shows that preemption does not require the cutting of a causal chain. Nevertheless, the most familiar sort of preemption does work by cutting. The causal chain from the preempting cause gets in first: it runs to completion, and the effect happens, while the

[2] For my former view, see the treatment of preemption in "Postscript E to 'Causation'," in my *Philosophical Papers*, Volume II (New York: Oxford, 1986), pp. 193-212.

Two points of terminology. Some say 'overdetermination' to cover all sorts of redundancy; I limit it to the symmetrical cases. Some say 'preemption' to cover only those asymmetrical cases which do involve cutting; I apply it to all asymmetrical cases.

[3] The discovery of trumping is due to Jonathan Schaffer, "Trumping Preemption," this JOURNAL, this issue, pp. 165-81. The example of the soldiers is due to Bas van Fraassen.

chain from the preempted alternative is still on its way. The preempted chain is cut. The effect itself is what prevents its final steps.

Billy and Suzy throw rocks at a bottle. Suzy throws first, or maybe she throws harder. Her rock arrives first. The bottle shatters. When Billy's rock gets to where the bottle used to be, there is nothing there but flying shards of glass. Without Suzy's throw, the impact of Billy's rock on the intact bottle would have been one of the final steps in the causal chain from Billy's throw to the shattering of the bottle. But, thanks to Suzy's preempting throw, that impact never happens.

I used to call such cases as this *late preemption*; in hindsight, *late cutting* is a better name. I meant to contrast them with *early preemption* (or better, *early cutting*): easy cases in which we have, if not direct counterfactual dependence of the effect upon the preempting cause, at least *stepwise dependence*. The effect depends counterfactually upon some intermediate event, which in turn depends upon the preempting cause. (Or there might be more steps.) In such cases, we get the right answer if we take causation to be the ancestral of dependence.

There is a small industry devoted to solving the preemption problem under the presupposition that all preemption works by cutting. However well such solutions may (or may not) work in the cases for which they were made, they are not general solutions because they cannot handle trumping.

IV. QUASI-DEPENDENCE REJECTED

I used to think that late cutting could be handled by appealing to the intuitive idea that causation is an intrinsic relation between events. (Except insofar as being subject to such-and-such laws of nature is an extrinsic matter, as I believe it to be.) Take another case, actual or possible, which is intrinsically just like the case of Suzy's throwing her rock at the bottle (and which occurs under the same laws), but in which Billy and his rock are entirely absent. In this comparison case, we have a causal chain from Suzy's throw to the shattering which *does* exhibit counterfactual dependence and which is, near enough, an intrinsic duplicate of the actual chain from Suzy's throw with Billy present. I thought: if being a causal chain is an intrinsic matter, then both or neither of these two chains must be causal; but the comparison chain, which exhibits dependence, surely is a causal chain; so the actual chain, even though thanks to Billy it does not exhibit dependence, must be a causal chain, too. I said that the actual chain exhibited *quasi-dependence*: it qualified as causal by courtesy, in virtue of its intrinsic resemblance to the causal chain in the comparison case (*op. cit.*, pp. 205-07).

Quasi-dependence was a bad idea, for several reasons. Here, one reason will suffice. Quasi-dependence gives the wrong answer in cases of trumping preemption. The trumped causal chain runs to completion; therefore, it is an intrinsic duplicate (near enough) of an *un*trumped causal chain in a comparison case (under the same laws) that exhibits counterfactual dependence. This shows at least that the invocation of quasi-dependence fails in possible worlds where trumping takes place. But worse, it may mean that the intrinsic character of causation is an over-hasty generalization even about the causation that happens in our own world. It may be, for all we know, that our case of the soldiers obeying the major is a trumping case that actually happens.

V. FRAGILITY CORRECTED

There is an obvious solution to cases of late cutting. Without Suzy's preempting rock, the bottle would still have shattered, thanks to Billy's preempted rock; but this would have been a *different* shattering. It would, for instance, have happened a little later. The effect that actually occurred depended on Suzy's throw, but not on Billy's.

Sometimes this solution is just right, and nothing more need be said. Suppose it were alleged that since we are all mortal, there is no such thing as a cause of death. Without the hanging that allegedly caused the death of Ned Kelly, for instance, he would sooner or later have died anyway. Yes. But he would have died a different death, and the event that actually was Kelly's death would never have occurred.

The case of Suzy's preempting throw is different, however. It is not just that without it the bottle would have shattered somehow, sooner or later. Without it, the bottle would have shattered at almost the same time that it actually did shatter, in almost the same way that it actually did. Yet we are usually quite happy to say that an event might have been slightly delayed, and that it might have differed somewhat in this or that one of its contingent aspects. So if we say that the shattering of the bottle was caused by Suzy's throw, because without it that very shattering would not have occurred, we are thinking that it would take only a very slight difference to destroy that event altogether, and put a different substitute event in its place. We are supposing the shattering to be modally *fragile*. This is not something we would normally suppose. We have no business first saying as usual that the very same event might have been significantly delayed and changed, and then turning around and saying that it is caused by an event without which it would have been ever so slightly delayed and changed, and then saying that this is because it

takes only a very slight delay or change to turn it into a different event altogether.

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. We have not made up our minds; and if we presuppose sometimes one answer and sometimes another, we are entirely within our linguistic rights. This is itself a big problem for a counterfactual analysis of causation, quite apart from the problem of preemption.[4]

At least, it is a problem so long as we focus on whether-whether counterfactual dependence. But there are other kinds of dependence. There is, for instance, when-on-whether dependence: *when* one event occurs depends counterfactually on *whether* another event occurs. And that is only the beginning. But even this beginning is enough to rehabilitate the obvious solution to late preemption, at least in very many commonplace cases. Let us by all means agree that Suzy's throw caused the shattering of the bottle because, without her throw, the shattering would have been slightly delayed. But let us not go on to say that, if it had been slightly delayed, that would have turned it into a different event altogether. Let us rather say that Suzy's throw caused the shattering in virtue of when-on-whether dependence, since without her the shattering would not have occurred exactly when it actually did.

L. A. Paul[5] has proposed an emended counterfactual analysis of causal dependence: event E depends causally on a distinct actual event C if and only if, "if C had not occurred, then E would not have occurred at all *or would have occurred later than the time that it actually did occur*" (*ibid.*, p. 193). This proposal does not abandon the strategy of fragility, but corrects it. Instead of supposing that the event itself is fragile—which would fly in the face of much of our ordinary talk—we instead take a tailor-made fragile proposition about that event and its time. The negation of that fragile proposition is the consequent of our causal counterfactual. Now we get the right answer to our commonplace cases of late cutting. Suzy's throw hastens the shattering, Billy's does not. So Suzy's throw causes the shattering, Billy's does not.

[4] It is a problem that is seldom noted; however, see Jonathan Bennett, *Events and Their Names* (Indianapolis: Hackett, 1988), *passim.*

[5] "Keeping Track of the Time: Emending the Counterfactual Analysis of Causation," *Analysis*, LVIII (1998): 191-98.

If we stopped here, we would be building into our analysis an asymmetry between hasteners and delayers. We would be saying that an event without which the same effect would have happened later is a cause, but an event without which the same effect would have happened earlier is not.[6] So we should not stop here. We should admit delayers as causes, even when the delayed event is the very same event that would otherwise have happened earlier—or at least, to acknowledge our indecision about such questions, not clearly *not* the same event.

We are often ambivalent about the status of delayers. Perhaps that is because a delayer often causes a later version of the event by preventing an earlier version which, had it happened, would have prevented the later version. Then, if we ask whether the delayer prevented the event or caused it, and we overlook the possibility that it might have done both, we have to say 'prevented'.[7]

To restore symmetry between hastening and delaying, we need only replace the words 'or would have occurred later than the time that it actually did occur' by the words 'or would have occurred at a time different from the time that it actually did occur'. I favor this further emendation. (As does Paul.) But I think we should go further still. What is so special about time? When we thought that without the actual causes of his death, Ned Kelly would have died a different death, we were thinking not just that he would have died at a different time, but also that he would have died in a different manner. According to the uncorrected strategy of fragility, a difference either of time or of manner would suffice to turn the effect into a numerically different event. And if, imitating Paul's correction, we relocate the fragility not in the event itself but rather in a tailor-made proposition about that event, that will be a proposition about how and when and whether the effect occurs. We could further emend our analysis to require dependence of how and when and whether upon whether: without *C*, *E* would not have occurred at all, or would have occurred at a time different from the time that it actually did occur, or would have occurred in a manner different from the manner in which it actually did occur.

This formulation still distinguishes the case that event *E* occurs differently from the case that *E* does not occur at all. The distinction has been made not to matter, but we are still presupposing that

[6] For advocacy of just such an "asymmetry fact," see Bennett, "Event Causation: The Counterfactual Analysis," *Philosophical Perspectives*, I (1987): 367-86; for reconsideration and rejection of it, see Bennett, *Events and Their Names*, pp. 69-72.

[7] See Penelope Mackie, "Causing, Delaying, and Hastening: Do Rains Cause Fires?" *Mind*, CI (1992): 483-500; and on double prevention generally, see Ned Hall, "Causation and the Price of Transitivity," this JOURNAL, this issue, pp. 198-222.

there is a distinction. If we are as indecisive about such questions as I think we are, it would be better to avoid that presupposition.

Let an *alteration* of event *E* be either a very fragile version of *E* or else a very fragile alternative event that is similar to *E*, but numerically different from *E*. If you think *E* is itself very fragile, you will think that all its alterations (except for the actual alteration) are alternatives. If you think *E* is not at all fragile, you will think that all its alterations are versions. You might think that some are alternatives and others are versions. Or you might refuse to have any opinion one way or the other, and that is the policy I favor. Now we may return to whether-whether dependence, but with alterations of the effect put in place of the event itself: without *C*, the alteration of *E* which actually did occur would not have occurred. However indecisive we may be about how fragile an event itself is, its actual alteration is by definition fragile.

Now we say that Suzy's throw caused the shattering of the bottle and Billy's preempted throw did not because, without Suzy's throw, the alteration of the shattering which actually did occur would not have occurred, and a different alteration would have occurred instead. Here, we are considering not only the slight delay before Billy's rock arrived but also any differences to the shattering that might have been made because Billy's rock differs from Suzy's in its mass, its shape, its velocity, its spin, and its aim point.

VI. SPURIOUS CAUSATION

We have dealt with one objection against the fragility strategy: that it conflicts with what we normally think about the conditions of occurrence of events. But there is a second objection, and it applies as much to the corrected strategy as to the strategy in its original form. All manner of irrelevant things that we would not ordinarily count among the causes of the effect can be expected to make some slight difference to its time and manner. I once gave this example: if poison enters the bloodstream more slowly when taken on a full stomach, then the victim's death, taken to be fragile—we might better say, the actual alteration of the victim's death—depends not only on the poison but also on his dinner (*op. cit.*, pp. 198-99).[8] If we heed still smaller differences, almost everything that precedes an event will be counted among its causes. By the law of universal gravitation, a distant planet makes some minute difference to the trajectory of Suzy's rock, thereby making a tiny difference to the shattering of the bottle. So by adopting the fragility strategy, in whichever form, we open the gate to a flood of spurious causes.

[8] Here I am indebted to Ken Kress.

Among the spurious causes that should have been deemed irrelevant is Billy's rock, the preempted alternative. We wanted to say that (the actual alteration of) the shattering depended on Suzy's throw and not on Billy's, but that turns out to be not quite true.

Well, these differences made by spurious causes are negligible; so, surely, we are entitled to neglect them. Just as it is right to say that a box contains nothing when, strictly speaking, it contains a little dust, so likewise we are within our linguistic rights to say that Billy's throw made no difference to the shattering when, strictly speaking, its gravitational effects made an imperceptibly minute difference. And if for some strange reason we did attend to these negligible differences, would we not then put ourselves in an unusual context where it was right, not wrong, to count all the things that make negligible differences as joint causes of the effect?

That would be a sufficient reply, I think, but for the fact that sometimes the difference made by a preempting cause is also minute. Imagine that Suzy's throw precedes Billy's by only a very short time; and that the masses, shapes, velocities, spins, and aim points of the two rocks also differ very little. Then without Suzy's throw we might have had a difference equal to, or even less than, some of the differences made by causes we want to dismiss as spurious.

But even so, and even if Billy's rock makes a minute difference to the shattering by way of its gravitational effects on Suzy's rock, yet Suzy's throw may make much *more* of a difference to the effect than Billy's. The alteration that would have occurred without Suzy's throw, though not very different from the actual alteration, may differ from it in time and manner more than the alteration that would have occurred without Billy's. That would be enough to break the symmetry between Suzy and Billy, and to account for our judgment that Suzy's throw and not Billy's causes the shattering. We speak of the asymmetry as if it were all-or-nothing, when really it is a big difference of degree, but, surely, such linguistic laxity is as commonplace as it is blameless.

If, on the other hand, Billy's throw does somehow make roughly as much difference to the effect as Suzy's, that is a good reason to say that it is not after all a mere preempted alternative. Rather, it is a joint cause of the shattering. In this case, too, we get the right answer.

VII. ALTERATIONS OF THE CAUSE

Because we are so indecisive about the distinction between alterations that are different versions of the same event and alterations that are different but similar events, we ought to make sure that this distinction bears no weight in our analyses. So far, we are obeying

that maxim only one-sidedly. The distinction does not matter when applied to the effect, but it still matters when applied to the cause. What it means to suppose counterfactually that C does not occur depends on where we draw the line between C's not occurring and C's occurring differently in time and manner.

That makes a problem. 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.[9] When asked to suppose counterfactually that C does not occur, we do not really look for the very closest possible world where C's 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. One repair would be to rewrite our counterfactual analysis, or add a gloss on its interpretation, in order to make this explicit (*op. cit.*, p. 211).

But there is another remedy. We could look at a range of alterations of C, not just one. As on the side of effects, we need not say which, if any, of these are versions of C and which, if any, are alternatives to C. These alterations may include some in which C is completely excised, but we need not require this. They may include some that are almost but not quite C, but we say nothing that restricts us to the closest possible alterations. Then we look at the pattern of counterfactual dependence of alterations of the effect upon alterations of the cause. Where C and E are distinct actual events, let us say that C *influences* E if and only if there is a substantial range C_1, C_2...of different not-too-distant alterations of C (including the actual alteration of C) and there is a range E_1, E_2...of alterations of E, at least some of which differ, such that if C_1 had occurred, E_1 would have occurred, and if C_2 had occurred, E_2 would have occurred, and so on. Thus we have a pattern of dependence of how, when, and whether upon how, when, and whether.

Influence admits of degree in a rough and multidimensional way. How many, and how varied, are the C_is? How distant are the C_is from one another; and, especially, how distant are the rest of them

[9] See Bennett, "Event Causation: The Counterfactual Analysis," pp. 369-70. His point here is independent of his defense, elsewhere in that article, of a hastener-delayer asymmetry.

from the actual alteration of C? How much do the E_is differ from one another? Plainly, there are many ways in which something can be more of a cause of some effect than something else is, even if it is not an all-or-nothing difference of influence versus no influence.

Now we are in a better position than before to say that Suzy's throw is much more of a cause of the bottle's shattering than Billy's. Even if the throws are so much alike that removing Suzy's throw altogether would make little difference to the shattering, it is still true that altering Suzy's throw slightly while holding Billy's fixed would make a lot of difference to the shattering, but altering Billy's throw slightly while holding Suzy's fixed would not. Take an alteration in which Suzy's rock is heavier, or she throws a little sooner, or she aims at the neck of the bottle instead of the side. The shattering changes correspondingly. Make just the same alterations to Billy's preempted throw, and the shattering is (near enough) unchanged.

Thanks to this latest emendation of the counterfactual analysis, cases of trumping are covered along with commonplace preemption. Sergeant and major both shout 'Advance!'. The soldiers advance. Altering the major's command while holding the sergeant's fixed, the soldiers' response would have been correspondingly altered. If the major had said 'Take cover!' they would have taken cover, if he had said 'Retreat!' they would have retreated, and so on. Altering the sergeant's command while holding the major's fixed, on the other hand, would have made (near enough) no difference at all. If we look only at the whether-whether dependence of the soldiers' response on the actual commands of the two officers, we miss exactly the sort of counterfactual dependence that breaks the symmetry between the two.[10]

VIII. TRANSITIVITY OF CAUSATION

Causation, I previously said, is the ancestral of causal dependence. Event C causes event E if and only if there is a chain of dependencies running from C to E. Is it still necessary to take the ancestral, now that our definition of causal dependence has evolved from simple whether-whether dependence to a pattern of influence? Does our improved definition of dependence allow us just to identify causation with causal dependence? No. Influence is not invariably transitive. If we want to ensure that causation is invariably transitive, we still have to take an ancestral.

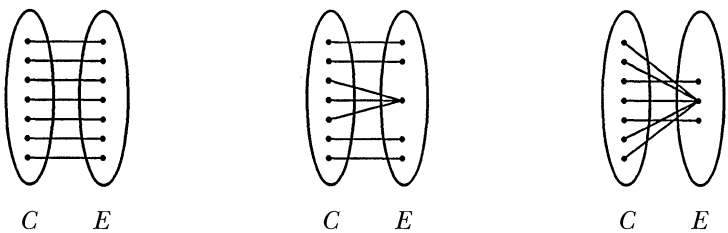
You might think that intransitivities of influence could arise from intransitivities of counterfactuals themselves. We know it can be true that, if P, it would be that Q, and true also that, if Q, it would be that

[10] Here I am indebted to Hall.

R, yet false that, if *P*, it would be that *R*. But that is not the problem. Although counterfactual transitivity itself is fallacious, a closely related inference pattern is valid: from the premise that, if *P*, it would be that *Q*, and the premise that, if both *P* and *Q*, it would be that *R*, it does follow that, if *P*, it would be that *R*. Let the counterfactual from C_i to D_i be part of a pattern of influence of *C* upon *D*; let the counterfactual from D_i to E_i be part of a pattern of influence of *D* upon *E*; then it would seem that, if both C_i and D_i, it would be that E_i so we do indeed have the counterfactual from C_i to E_i; and likewise for the other counterfactuals that constitute a pattern of influence of *C* on *E*.

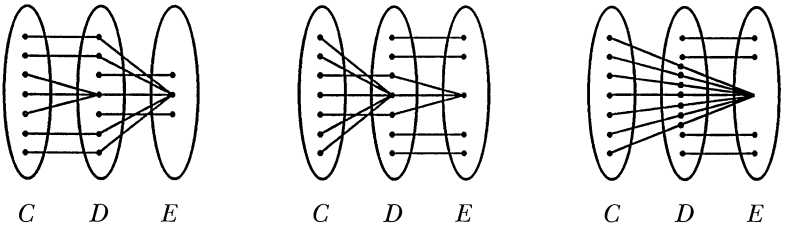
The real problem with transitivity is that a pattern of influence need not map *all* the not-too-distant alterations of *C* onto different alterations of *D*, or all the not-too-distant alterations of *D* onto different alterations of *E*. Transitivity of influence can fail because of a mismatch between the two patterns of influence.

Below I picture three possible patterns of influence of *C* on *E*. The first is nice and simple: it maps several alterations of *C* one-one onto alterations of *E*. But less nice patterns will still qualify. Let the actual alteration be at the center, and imagine that distance from the center somehow measures closeness to actuality. We might have a pattern that maps the outer alterations of *C* (second picture) or the inner alterations of *C* (third picture) one-one onto different alterations of *E*, but funnels the other alterations of *C* onto a single point.



Now, suppose *C* influences *D* by a pattern that funnels all the inner alterations onto a single point, while *D* influences *E* by a pattern that funnels all the outer alterations onto a single point (first picture below), or vice versa (second picture). Or we might have more complicated cases (third picture). In each case, the two patterns of influence which take us from *C* to *D* to *E* are mismatched: the values of the first pattern do not coincide with the arguments of the second. So *C* influences *D* and *D* influences *E*, but *C* does not influence

E. If we nevertheless want to say that *C* causes *E*, we have to take the ancestral and say that causation outruns direct influence.



How might such a case arise? Here is a famous example from Harry Frankfurt.[11] The neuroscientist knows just how she wants Jones to behave. She hopes that Jones, left to himself, will behave just as she wants. By reading his brain, she can predict what he will do if left to himself. She predicts that he will do just what she desires, so she does nothing else. But if instead she had read that he would stray from the desired path, she would have taken control and manipulated his brain directly so as to produce the desired behavior. The initial state of Jones's brain is a preempting cause of his behavior; the idle neuroscientist is a preempted backup. The moral of the story is that preemptive causation, without dependence, suffices to confer ownership and responsibility for one's actions.

Let *C* be Jones's initial brain state; let *E* be the desired behavior. Consider a time after the neuroscientist has read Jones's brain, but before she would have seized control if the reading had been different. Let *D* combine Jones's brain state at that time with the neuroscientist's decision not to intervene. We have a two-step chain of influence from *C* to *D* to *E*. But *C* does not influence *E*. Any alteration of Jones's initial brain state would have led to the desired behavior in the end, one way or the other.

The actual alteration of *C* is the one that leads to the desired behavior. The actual alteration of *E* consists of the desired behavior; the other alterations of *E* consist of different behavior. The actual alteration of *D* is the one that leads to the desired behavior, and that includes the neuroscientist's decision not to intervene. The "inner" alterations of *D* are those which would not lead to the desired behavior, but which include the neuroscientist's decision to intervene. The "outer" alterations of *D* are those which would not lead to the desired behavior, but which nevertheless include the neuroscientist's

[11] "Alternate Possibilities and Moral Responsibility," this JOURNAL, LXVI, 23 (December 4, 1969): 829-39.

decision not to intervene. These are arguments of the pattern from *D* to *E*, but not values of the pattern from *C* to *D*. The example illustrates the third sort of mismatch shown above. Transitivity of influence fails.

This is an easy case of early cutting—just the sort of case for which my strategy of taking the ancestral was originally made. If we had tried to make do without the ancestral, and to get by with influence alone, it would remain unsolved—provided that we insist, as of course we should, that with no intervention by the neuroscientist, Jones's initial brain state is indeed a cause of his behavior.

IX. TRANSITIVITY DEFENDED

Some will say that, by making causation invariably transitive, our strategy of taking the ancestral makes more trouble than it cures. It collides with a flock of alleged counterexamples to the transitivity of causation. Thus, I have incurred an obligation to deal with these examples.

The counterexamples have a common structure. Imagine a conflict between Black and Red. (It may be a conflict between human adversaries, or between nations, or between gods striving for one or another outcome, or just between those forces of nature which conduce to one outcome versus those which conduce to another.) Black makes a move that, if not countered, would have advanced his cause. Red responds with an effective countermove, which gives Red the victory. Black's move causes Red's countermove, Red's countermove causes Red's victory.[12] But does Black's move cause Red's victory? Sometimes, it seems not.

My considered opinion is that Black's move does indeed cause Red's victory. Transitivity succeeds. But I admit to feeling some ambivalence. Insofar as I can summon up any inclination to accept the counterexamples, I think my inclination has three sources, all of them misguided.

First. In many of these cases Red's victory would have come sooner, or more directly, without Black's move. Black's move prevents Red's victory as well as causing it: it causes one version, but it prevents another. If we thought we had to choose, we would wrongly infer that since it is a preventer it cannot be a cause.

Second. Moves such as Black's are in general conducive to victory for Black, not for Red. If we mix up questions of what is generally conducive to what with questions of what caused what in this particu-

[12] Examples of this form appear in Bennett, "Event Causation: The Counterfactual Analysis," p. 373; Michael McDermott, "Redundant Causation," *British Journal for the Philosophy of Science*, XL (1995): 523-44; and Hall.

lar case, we may think it just a bit of good common sense to say that Black's moves advance Black's cause, not Red's.[13]

Third. We note that Black's move did not matter; Red would have won all the more easily without it. The effect does not depend on the cause. The idea that causation requires whether-whether dependence may retain some grip on us. But if you *ever* accept preemptive causation, you must have learned to resist that grip. Why yield to it now? It is true that Black's move did not matter. But that is because the choice Black faced (whether he knew it or not) was whether to have his defeat caused in one way or in another. Either way, Black's defeat is *caused.*

In rejecting the counterexamples, and accepting that Black's move is a cause of Red's victory, I think I am doing what historians do. They trace causal chains, and, without more ado, they conclude that what comes at the end of the chain was caused by what went before. If they did not, they could say little about historical causation; because, over intervals of any length, historical counterfactuals become so very speculative that nothing much can be known about the dependence of any event on its causal ancestors. And every historian knows that actions often have unintended and unwanted consequences. It would be perfectly ordinary for a move by Black to backfire disastrously.[14]

X. CAUSATION BY ABSENCES

Alterations, I said, are very fragile events. That was not quite right: some of them are absences. Absences can be causes, as when an absence of food causes hunger; they can be effects, as when a vaccination prevents one from catching a disease; and they can be among the unactualized alterations of a cause or effect which figure in a pattern of influence.

Absences are not events. They are not *anything*: where an absence is, there is nothing relevant there at all. Absences are bogus entities. Yet the proposition that an absence occurs is not bogus. It is a perfectly good negative existential proposition. And it is by way of just such propositions, and only by way of such propositions, that ab-

[13] Compare Lawrence Lombard, "Causes, Enablers, and the Counterfactual Analysis," *Philosophical Studies,* LIX (1990): 195-211, here p. 197.

[14] I have assumed so far that the Black-Red examples are genuine test cases: we really do have an event C that causes an event D that, in turn, causes an event E. But unless the examples are carefully formulated, perhaps with the aid of somewhat artificial stipulations, that may not be so. It may rather be that C causes D_1 and D_2 causes E; and D_1 and D_2 are different, even though perhaps we may refer to them by the same name. D_1 and D_2 might be two different spatiotemporal parts, or two different logical parts, of the same event. See Paul, "Aspect Causation," this JOURNAL, this issue, pp. 235-56.

sences enter into patterns of counterfactual dependence. Therefore, it is safe to say with the vulgar that there are such entities as absences, even though we know better.

One reason for an aversion to causation by absences is that, if there is any of it at all, there is a lot of it—far more of it than we would normally want to mention. At this very moment, we are being kept alive by an absence of nerve gas in the air we are breathing. The foe of causation by absences owes us an explanation of why we sometimes do say that an absence caused something. The friend of causation by absences owes us an explanation of why we sometimes refuse to say that an absence caused something, even when we have just the right pattern of dependence. I think the friend is much better able to pay his debt than the foe is to pay his. There are ever so many reasons why it might be inappropriate to say something true. It might be irrelevant to the conversation, it might convey a false hint, it might be known already to all concerned....

Of course, such reasons for refusing to say what is true are not confined to causation by absences. "Counterfactual analysis of causation?—Yeah, yeah, my birth is a cause of my death!" said the scoffer. His birth is indeed a cause of his death; but it is understandable that we seldom say so. The counterfactual dependence of his death on his birth is just too obvious to be worth mentioning.[15]

It does not make sense for two distinct absences to differ slightly in detail. When we have an absence, there is nothing (relevant) there at all, and that is that. So when an absence is caused, we would expect a pattern of influence which exhibits funneling to an unusual degree. To illustrate funneling, we can imagine a device that works in an extraordinarily precise all-or-nothing fashion; or a neuroscientist, or some other marvelous being, able to exert extraordinarily precise and complete control; or we can just imagine a perfectly ordinary case of prevention. If we then follow that with the funneling that comes from the presence of a preempted backup, we may well end up with mismatched patterns of influence in which transitivity of influence fails. Small wonder, then, that cases of preemptive prevention have appeared along with the Black-Red examples in the debate over transitivity of causation. I say again that at worst we have causation without direct influence. I trace a chain; I take the ancestral; I say that when a preempted preventer causes an absence which, in turn, causes some further event or absence, then the preempted preventer is a cause of that further event or absence.

[15] See H.P. Grice, "Logic and Conversation," in his *Studies in the Way of Words* (Cambridge: Harvard, 1989), pp. 22-40.

Part of what makes these cases hard, however, is doubt about whether the absence does cause anything further. The fielder catches the ball; he causes its absence from the place just beyond his hand. But a little further along its path there is a wall—a high, broad, thick, sturdy wall.[16] Further along still is a window. Does the fielder cause the window to remain unbroken? We are ambivalent. We have *C*, the catch. We have *D*, the absence of the ball from the place just beyond the fielder's hand. We have *E*, the absence of the impact of the ball on the window, or the nonbreaking of the window. Certainly, we have a pattern of influence of *C* upon *D*. Whether we have influence of *D* upon *E* is doubtful. There are alterations of *D* in which not only is the ball present beyond the fielder's hand, but also it is on a trajectory that would take it over the high wall and down again, or it is moving with energy enough to break through, or.... Some of these alterations of *D* would indeed have led to alterations of *E*. But are they relevant, "not-too-distant," alterations of *D*? We may be in a mood to think so, or we may be in a mood to think not. If we are in a mood to think them relevant, we should conclude that *D* causes *E*, and, by transitivity, *C* also causes *E*. Whereas if we are in a mood to think them not relevant, we should conclude that *D* does not cause *E*, *C* does not cause *E*, and issues of transitivity do not arise.

DAVID LEWIS

Princeton University

[16] The example of the fielder comes from McDermott. The suggestion that our wavering intuitions are governed by how far-fetched we find the possibility of the ball's getting past the wall comes from John Collins, "Preemptive Prevention," this JOURNAL, this issue, pp. 223-34.